Defence Research and Development Canada
Recherche et développement pour la défense Canada

Bayesian Inversion of Concentration Data for an Unknown Number of Contaminant Sources

E. Yee
DRDC Suffield

Technical Report
DRDC Suffield TR 2007-085
June 2007

Canada

Bayesian Inversion of Concentration Data for an Unknown Number of Contaminant Sources

Eugene Yee
Defence R&D Canada – Suffield

Defence R&D Canada – Suffield
Technical Report
DRDC Suffield TR 2007–085
June 2007

Author

Original signed by Eugene Yee

Eugene Yee

Approved by

Original signed by Joan Armour

Dr L. Nagata
Head, Chemical and Biological Defence Section

Approved for release by

Original signed by Dr P. D'Agostino

Dr P. D'Agostino
Chair, DRDC Suffield DRP

Abstract

In this paper, we address the inverse problem of source reconstruction for the difficult case of multiple sources when the number of sources is unknown *a priori*. The problem is solved using a Bayesian probabilistic inferential framework in which Bayesian probability theory is used to derive the posterior probability density function for the number of sources and for the parameters (e.g., location, emission rate, release duration) that characterize each source. A mapping (or, source–receptor relationship) that relates a multiple source distribution to the concentration data measured by the array of sensors is formulated based on a forward-time Lagrangian stochastic model. A computationally efficient methodology for determination of the likelihood function for the problem, based on an adjoint representation of the source–receptor relationship and realized in terms of a backward-time Lagrangian stochastic model, is described. An efficient computational algorithm based on a reversible jump Markov chain Monte Carlo (MCMC) method is formulated and implemented to draw samples from the posterior density function of the source parameters. This methodology allows the MCMC method to jump between the hypothesis spaces corresponding to different numbers of sources in the source distribution and, thereby, allows a sample from the full joint posterior distribution of the number of sources and source parameters to be obtained. The proposed methodology for source reconstruction is tested using synthetic concentration data generated for cases involving two and three unknown sources.

Résumé

Dans cet article, nous abordons le problème inverse de reconstruction d'une source, concernant le cas difficile de sources multiples, quand le nombre de source est à priori inconnu. Le problème est résolu au moyen d'un concept inférenciel de probabilité bayésienne avec lequel on utilise la théorie de probabilité bayésienne pour dériver la densité de probabilité postérieure pour le nombre de sources et les paramétres (ex. : endroit, taux d'émission, durée d'émission) qui caractérisent chaque source. Un mappage (ou une relation source–récepteur), faisant la relation entre la répartition des sources multiples et les données de concentration mesurées par le réseau de capteurs, est formulé, basé sur le modèle stochastique Langrangian futuriste (forward-time) On y décrit une méthodologie computationnelle efficiente visant à déterminer la fonction de vraisemblance pour le problème, basée sur une représentation adjointe de la relation source–récepteur et réalisée en termes d'un modèle stochastique Langrangian passéiste(backward-time). Un algorithme computationnel efficient basé sur une méthode Monte Carlo-chaîne de Markov (MCMC) de saut réversible a été formulé et implémenté pour tirer des échantillons provenant de la densité de probabilité postérieure des paramétres de la source. Cette méthodologie permet à la méthode MCMC de sauter entre les espaces d'hypothèse correspondant aux nombres de sources différentes dans la répartition des sources et, par conséquent, permet d'obtenir un échantillon tiré de la répartition postérieure des sources intégralement conjointes du nombre de sources et des paramètres de ces sources. Cette méthodologie proposée pour la reconstruction des sources est testée au moyen de données de concentrations synthétiques générées pour des cas ayant deux ou trois sources inconnues.

Executive summary

Introduction: An increasingly capable sensing technology for concentration measurements of chemical, biological and radiological (CBR) agents released into the atmosphere, either accidentally or deliberately, has fostered interest in exploiting this information for detection, identification and reconstruction of agent sources responsible for the observed concentration. A critical capability gap in current emergency and retrospective management efforts, directed at terrorist incidents involving the covert release of a CBR agent in a densely populated urban center, is the determination of the number of sources and for each of these sources the location and amount of agent released (source term estimation problem), following event detection by an array of independent CBR sensors. In order to address this capability gap, an additional component for source reconstruction has been included in the Chemical Biological Radiological and Nuclear (CBRN) Research and Technology Initiative (CRTI) project 02-0093RD (Advanced Emergency Response System for CBRN Hazard Prediction and Assessment in the Urban Environment). This component has has been incorporated in CRTI 02-0093RD under the auspices of the Public Security Technical Program (PSTP) project "United States/Canada Collaborative Projects on Science and Technology in Urban Transport Modeling Related to Homeland Security" and funded under the CRTI Supplemental Funding Program for existing projects, with Defence R&D Canada – Suffield as the principal investigator.

Results: In this report, we address the problem of source term estimation for the difficult case of multiple sources when even the number of sources is unknown *a priori*. A new methodology was developed to solve this problem. This methodology provides simultaneous estimates for the number of sources and for the parameters (e.g., location, emission rate, release duration) that characterize each source. The proposed methodology for source reconstruction has been tested successfully using synthetic concentration data generated for cases involving two and three unknown agent sources.

Significance and Future Plans: The algorithm developed in this report may be used to interpret agent concentration measurements obtained from an array of sensors in order to estimate unknown source characteristics and quantify the uncertainties in this estimation. The technique may be used to infer the location, emission rate, and duration of putative agent release(s) and this information can be subsequently used to predict future agent transport in the atmospheric environment. In summary, the algorithm provides a methodology for the fusion of CBR sensor data with model predictions of agent dispersal in the atmosphere to provide a more complete situational awareness in the operational environment. Future plans include testing the proposed Bayesian inference algorithm for multiple sources using real concentration measurements obtained from a cooperative **FU**sing **S**ensor **I**nformation from **O**bserving **N**etworks (FUSION) Field Trial 2007 (FFT 07) that will be conducted at U.S. Army Dugway Proving Ground (DPG) in September 2007. This field experiment is conducted under Technical Panel 9 (TP-9) of The Technical Cooperation Program (TTCP) Chemical, Biological and Radiological Defense (CBD) Group.

Yee, E. (2007). Bayesian Inversion of Concentration Data for an Unknown Number of Contaminant Sources. (DRDC Suffield TR 2007–085). Defence R&D Canada – Suffield.

Sommaire

Introduction: Une technologie de détection de plus en plus efficace à mesurer la concentration d'agents chimiques, biologiques et radiologiques (CBR) libérés dans l'atmosphère, soit accidentellement soit délibérément, a suscité l'intérêt concernant l'exploitation de cette information pour la détection, l'identification et la reconstruction de sources d'agents ayant causé la concentration observée. Diminuer l'écart crucial entre la capacité des efforts à répondre aux urgences et celle de leur gestion rétrospective dirigée contre des incidents terroristes comprenant la dissémination secrète d'agent CBR dans un centre urbain densément peuplé, consiste à déterminer le nombre de sources et pour chacune de ces sources, l'endroit et la quantité d'agent disséminé (problème d'estimation en termes de sources) qui suit la détection d'un événement au moyen d'un réseau de capteurs CBR indépendants. Pour être en mesure de diminuer cet écart entre les capacités, une composante complémentaire de la reconstruction des sources a été inclue dans le Projet 02-0093RD de l'Initiative de recherche et de technologie chimique, biologique, radiologique et nucléaire CBRN (IRTC) (Système d'intervention d'urgence avancé pour la prédiction et l'évaluation des risques CBRN en milieu urbain). Cette composante a été incorporée dans le Projet 02-0093RD de l'IRTC sous l'égide du projet ≪ Projets communs entre le Canada et les États-Unis sur la science et la technologie de la modélisation des transports en commun reliés à la Sécurité interne ≫ du Programme technique de sécurité publique (PTST) qui est financé par le Programme de financement supplémentaire de l'IRTC, pour les projets existants, avec R&D pour la défense Canada – Suffield pour instigateur principal.

Résultats: Dans cet article, nous abordons le problème inverse de reconstruction d'une source, concernant le cas difficile de sources multiples, même quand le nombre de source est à priori inconnu. On a mis au point une nouvelle méthodologie pour résoudre ce problème. Cette méthodologie procure des estimations simultanées pour le nombre de sources et leurs paramètres (ex. : endroit, taux d'émission, durée d'émission) caractérisant chaque source. Les essais sur cette méthodologie proposée ont produit des résultats positifs pour la reconstruction des sources au moyen de données de concentrations synthétiques générées pour des cas ayant deux ou trois sources inconnues.

La Portée des Résultats et les Plans Futurs: L'algorithme mis au point dans ce rapport peut être utilisé pour interpréter les mesures de concentration d'agent obtenues d'un réseau de capteurs en vue d'estimer les caractéristiques d'une source inconnue et de quantifier les incertitudes contenues dans l'estimation. La technique peut être utilisée pour inférer l'endroit, le taux d'émission, la durée des disséminations de l'agent putatif et cette information peut être consècutivement utilisée pour prédire le transport futur de l'agent dans le milieu atmosphérique. En résumé, l'algorithme procure une méthodologie pour la fusion des données des capteurs CBR avec des prédictions d'un modèle de dispersion d'un agent dans l'atmosphère qui visent à procurer une connaissance plus complète de la situation du milieu opérationnel. Les plans futurs consistent à évaluer l'algorithme d'inférence bayésienne pour des sources multiples utilisant des mesures réelles de concentration obtenues en coopération de l'information fusionnée des capteurs du réseau météorologique (FUSION) des Essais sur le terrain 2007 (FFT 07) qui seront conduits sur le Polygone d'essai Dugway de l'arme américaine en septembre 2007. Cette expérience sur le terrain est conduite sous la direction du Panel technique 9 (PT9) du Programme de coopération technique (PCT) du groupe de

la défense chimique, biologique et radiologique.

Yee, E. (2007). Bayesian Inversion of Concentration Data for an Unknown Number of Contaminant Sources. (DRDC Suffield TR 2007–085). R & D pour la défense Canada – Suffield.

Table of contents

List of Figures

List of Tables

Introduction

The ever increasing power and sophistication of today's chemical, biological, and radiological (CBR) detection technologies are opening up a new realm for the detection and monitoring of releases of these toxic agents into the atmosphere. Rapid detection and early response to a release of a CBR agent could dramatically reduce the extent of human exposure (and associated morbidity and mortality of the incident) and minimize the cost of the subsequent clean up. In the event of a CBR incident, the assessment of the damage likely to be caused by the release is a problem of great importance. This assessment is usually undertaken using a predictive model for the mean transport and turbulent diffusion of the contaminant through the atmosphere, which in turn provides the information required to determine the temporal window and geographical extent of the hazard zone (or, toxic corridor) required in the formulation of an effective response. Unfortunately, an array of CBR sensors by itself is not sufficient for this task, owing to the fact that detection of a toxic agent plume by the sensor array only indicates that a release has occurred, but without knowing the characteristics of the source (which include things like the rate and duration of the release, source location, etc.), the prediction of the dispersion of the contaminant in the atmosphere cannot be made. Indeed, the principal impediment and critical capability gap in current emergency management efforts directed at terrorist incidents involving the covert release of a CBR agent into the atmosphere is the determination of the unknown source characteristics following event detection by a network of CBR sensors. This realization has provided the impetus for our recent research efforts directed towards a solution of the source reconstruction problem.

Although the problem of the forward prediction of the concentration field resulting from the dispersion of a contaminant released in a complex turbulent flow (with a known source) has received considerable attention from researchers, the "reverse" prediction of the source location and strength using a finite number of noisy concentration data obtained from an array of sensors has received considerably less attention, although the importance of the solution of this problem for a number of practical applications is obvious. Early work on inverse source determination focussed on the problem of source strength estimation whereby the location of the source was assumed to be known exactly *a priori*. For example, Hanna et al. [1] used three different types of dispersion models to estimate the emission rate (source strength) for the localized sources used in the Project Prairie Grass experiments. Gordon et al. [2] and Wilson and Shum [3] applied a forward-time Lagrangian stochastic trajectory model in conjunction with a mass balance technique to estimate the rate of ammonia volatilization from field plots (viz., area sources on the ground surface), while Flesch et al. [4] used a backward-time Lagrangian stochastic dispersion model to estimate the emission rate from a sustained surface area source over simple terrain. Kaharabata et al. [5] used an approximate solution to the advection–diffusion equation to determine the source strength from microplots over an open field, using single and multipoint measurements of the concentration downwind of the contaminated field plot. For longer-range dispersion, Robertson and Langner [6] estimated the emission rates for the European Tracer Experiment (Nodop et al. [7]) using a variational data assimilation procedure to minimize a cost function defined to be the sum of squared differences between the model solution and the concentration measurements over a 24-h assimilation period (window). A different approach was proposed by Siebert and Stohl [8], who attempted to reconstruct the source strength using Tikhonov regularization. Finally, Skiba [9] used an adjoint transport model to assess emission rates from various industrial plants in Mexico, whose locations were assumed to

be known *a priori*.

A number of researchers have focussed on the problem of recovering the location of an unknown source on the basis of available concentration measurements, given that the source strength is known *a priori*. Khapalov [10] formulated the source localization problem as a nonlinear identification problem based on the diffusion system modeling involving a distributed system of parabolic type. Another optimization approach was followed by Alpay and Shore [11] who described an "intelligent" brute-force method for source localization, which is based on the simulation of every possible scenario, with the likely source location being that which minimizes an appropriate norm between the model predictions of concentration for that source location and the measured concentration data. Matthes et al. [12] employed a two-step approach for source localization for the simple (albeit unrealistic) case of an emitted substance transported by homogeneous advection and isotropic diffusion. This method was designed to overcome convergence difficulties that have been experienced by various researchers using gradient-based optimization methods for formulation of the source localization problem.

A simultaneous estimation of the source location and strength was formulated by Alpay and Shore [11], who introduced the dual system method whereby the forward governing equation for the concentration is replaced with the adjoint equation for an adjoint state variable. A similar approach, in the context of dispersion on a global scale, was followed by Pudykiewicz [13], who inferred source parameters (location, emission rate, and time of release) by using the influence function obtained as a solution of the adjoint tracer transport equation to estimate the source location, in conjunction with the Lagrange duality relation to estimate the strength of the source. In the line of probabilistic approaches, a Bayesian inferential methodology in conjunction with the adjoint method (in which the transport and diffusion of matter in turbulent flows was solved using either partial or stochastic differential equations backward in time in an Eulerian or Lagrangian description of turbulent dispersion, respectively) was proposed by Yee [14], Yee [15], Keats et al. [16], and Yee [17] for the simultaneous determination of all source parameters for dispersion in complex environments (e.g., urban dispersion, long-range dispersion, etc.).

In all previous studies cited here, with the exception of Yee [17], the problem of identification of source parameters was restricted to one source. The determination of the characteristics of multiple localized sources was briefly addressed in Yee [17], but the approach used there assumed that the number of sources was known *a priori*. The problem of reconstruction of an unknown number of sources using a finite number of noisy concentration measurements obtained from an array of sensors is of great interest for current emergency and retrospective management efforts directed at terrorist incidents involving the release of CBR agents in a densely populated urban center. In this application, the array of CBR sensors measures the superposition of concentrations arising from multiple interfering sources in which both number of sources and the characteristics (e.g., location, emission rate) of each source are unknown *a priori*. The determination of the number of sources is a difficult problem and involves some form of model selection. In this paper, we formulate a Bayesian inferential scheme for the joint determination of the number of contaminant sources and the parameters for each source, given a finite number of concentration measurements made by an array of sensors. The Bayesian formulation of a solution for this problem provides a computational challenge owing to the fact that the resulting integrals of the posterior distribution of the source parameters over the hypothesis space are not analytically tractable, and stan-

dard numerical methods for integration cannot be applied to give accurate results due to the potential high dimensionality of the hypothesis space. To overcome this problem, our proposed procedure uses a reversible–jump Markov chain Monte Carlo method to estimate the number of unknown sources and the relevant parameters (e.g., emission rate, source location, etc.) for each source.

Concentration data model: source–receptor relationship

Here we present our model for the mean concentration "seen" by a sensor and the actual concentration measurements made by a sensor, the latter of which must necessarily include several sources of "noise".

Model for mean concentration

To solve the source reconstruction problem using Bayesian probability theory, it is necessary to relate the hypotheses of interest concerning the source distribution to the available concentration data measured by the array of sensors. This is a source–receptor relationship that describes the transport of a substance (e.g., CBR material) through the atmosphere after it has been released from a known source distribution and, as such, provides a prediction of the average concentration of the substance in a small volume centered at any given spatial location during any given time interval. With reference to the source–receptor relationship, the description of the turbulent transport of a "marked" fluid element (viz., a fluid element marked by the contaminant substance as it moves through any portion of the source distribution) in a Largrangian reference frame is perhaps the most natural and simplest description of this physical phenomenon. This involves modeling dispersion through the random walks for "marked" fluid elements.

Let $p_L(\mathbf{x}, t|\mathbf{x}', t')d\mathbf{x}'$ be the probability that a fluid element, displaced by the background flow field from within an infinitessimal volume $d\mathbf{x}'$ of the point $\mathbf{x}'$ (at time t'), passes through the point $\mathbf{x}$ (at time t where $t > t'$) in a statistical ensemble of turbulent flows. If $S(\mathbf{x}', t')$ is the source density function at the point $\mathbf{x}'$ (time t'), the concentration C at time t and location $\mathbf{x}$ of the contaminant released within $\mathcal{D}$ (space occupied by the fluid) up to this time is given as follows:

$$C(\mathbf{x}, t) = \int_{\mathcal{D}} \int_{-\infty}^{t} p_L(\mathbf{x}, t|\mathbf{x}', t')S(\mathbf{x}', t')\, dt'\, d\mathbf{x}'. \tag{1}$$

The mean concentration $\overline{C}(\mathbf{x}_r, t_r)$ "seen" by a sensor at receptor location $\mathbf{x}_r$ and time t_r corresponds to an average of $C(\mathbf{x}, t)$ over the spatial volume and averaging time of the sensor, so

$$\overline{C}(\mathbf{x}_r, t_r) = \int_0^T dt \int_{\mathcal{D}} d\mathbf{x} C(\mathbf{x}, t)h(\mathbf{x}, t|\mathbf{x}_r, t_r) \equiv \langle C, h\rangle(\mathbf{x}_r, t_r). \tag{2}$$

Here, $h(\mathbf{x}, t|\mathbf{x}_r, t_r)$ is the spatial-temporal filtering (or, response) function (of $\mathbf{x}$ and t) for the concentration measurement made by the sensor at location $\mathbf{x}_r$ and at time t_r; and, $\mathcal{D} \times [0, T]$ is the space-time volume (of the sensor) over which the averaging is taken. Note

that in Equation (2), we expressed the mean concentration $\overline{C}(\mathbf{x}_r, t_r)$ "seen" by a sensor as the inner (or scalar) product $\langle C, h \rangle$ of the concentration C and sensor response function h.

In complex turbulent flows (e.g., atmospheric flow), there are no known analytical solutions for the transition probability density function (PDF) $p_L(\mathbf{x}, t|\mathbf{x}', t')$ required for the determination of the concentration of tracer [cf. Equation (1)]. Because the exact form of $p_L(\mathbf{x}, t|\mathbf{x}', t')$ is not known, we need to construct some approximation for this PDF. In the context of Lagrangian stochastic trajectory simulation methods, the transition PDF for a fluid element at position $\mathbf{x}$ at time t, which at the earlier time t' was at position $\mathbf{x}'$, is determined as the solution to the following Ito stochastic differential equation (see, Thomson [18]):

$$\begin{aligned} d\mathbf{X}(t) &= \mathbf{U}(t)\, dt, \\ d\mathbf{U}(t) &= \mathbf{a}(\mathbf{X}(t), \mathbf{U}(t), t)\, dt + \big(C_0 \epsilon(\mathbf{X}(t), t)\big)^{1/2}\, d\mathbf{W}(t), \end{aligned} \tag{3}$$

where $\mathbf{X}(t) \equiv (X_i(t)) = (X_1(t), X_2(t), X_3(t))$ and $\mathbf{U}(t) \equiv (U_i(t)) = (U_1(t), U_2(t), U_3(t))$ are the (Lagrangian) position and velocity, respectively, of a "marked" fluid element (or, particle) at time t (marked by the source as the fluid element passes through it at some earlier time t'), so $(\mathbf{X}, \mathbf{U})$ determines the state of the fluid particle at any time t after its intial release from the source distribution S. In Equation (3), C_0 is the Kolmogorov universal constant (associated with the Kolmogorov similarity hypothesis for the form of the Lagrangian velocity structure function in the inertial subrange); ϵ is the mean dissipation rate of turbulence kinetic energy; $d\mathbf{W}(t) \equiv (dW_i(t)) = (dW_1(t), dW_2(t), dW_3(t))$ are the increments of a vector-valued (three-dimensional) Wiener process (i.e., they have a Gaussian distribution with zero mean and variance dt (infinitessimal time step), and non-overlapping increments are statistically independent); and $\mathbf{a} \equiv (a_i) = (a_1, a_2, a_3)$ is the drift coefficient vector.

The model is fixed by the choice of the drift coefficients a_i $(i = 1, 2, 3)$. The most important criterion for the determination of a_i is the well-mixed criterion [18] which states that an initially well-mixed distribution of particles should remain so. One possible model for a_i that satisfies the well-mixed criterion was given by Thomson [18] for the case of Gaussian turbulence in three dimensions in which the (unconditional) PDF P_E of the velocity fluid elements (both marked and unmarked) has the form:

$$P_E(\mathbf{u}) = \frac{1}{(2\pi)^{3/2}\big(\det \Gamma_{ij}\big)^{1/2}} \exp\left(-\frac{1}{2}\big(u_i - \overline{U}_i\big)\Gamma_{ij}^{-1}\big(u_j - \overline{U}_j\big)\right), \tag{4}$$

where $\overline{U}_i$ is the mean Eulerian velocity and $\Gamma_{ij} \equiv \overline{(u_i - \overline{U}_i)(u_j - \overline{U}_j)}$ is the Reynolds stress tensor for the underlying turbulent flow (here, tensor notation with the Einstein summation convention is used and an overline is used to denote an ensemble average). It should be noted that there is an implicit dependence of P_E on $\mathbf{x}$ and t through its dependence on $\overline{U}_i \equiv \overline{U}_i(\mathbf{x}, t)$ and $\Gamma_{ij} \equiv \Gamma_{ij}(\mathbf{x}, t)$. In this context, Thomson's solution for a_i that is consistent with the well-mixed criterion assumes the following form:

$$a_i = -\frac{1}{2}\big(C_0 \epsilon\big)\Gamma_{ik}^{-1}\big(u_k - \overline{U}_k\big) + \frac{\phi_i}{P_E}, \tag{5a}$$

where

$$\frac{\phi_i}{P_E} \equiv \frac{1}{2}\frac{\partial \Gamma_{il}}{\partial x_l} + \frac{\partial \overline{U}_i}{\partial t} + \overline{U}_l \frac{\partial \overline{U}_i}{\partial x_l} + \left(\frac{1}{2}\Gamma_{lj}^{-1} \left[\frac{\partial \Gamma_{il}}{\partial t} + \overline{U}_m \frac{\partial \Gamma_{il}}{\partial x_m} \right] + \frac{\partial \overline{U}_i}{\partial x_j} \right) \left(u_j - \overline{U}_j \right) + \frac{1}{2}\Gamma_{lj}^{-1} \frac{\partial \Gamma_{il}}{\partial x_k} \left(u_j - \overline{U}_j \right) \left(u_k - \overline{U}_k \right). \tag{5b}$$

This model is general enough to allow for almost any flow complexity.

Equation (3), in conjunction with Equations (5a) and (5b), is a forward-time Lagrangian trajectory simulation model. In this source-oriented approach, "marked" fluid elements with initial space–time coordinates $(\mathbf{x}', t')$ are sampled from a space–time density function proportional to the (prescribed) source distribution $S(\mathbf{x}', t')$. The forward Lagrangian trajectories $(t > t')$ of these "marked" fluid elements, which emanate from the source and move toward the receptor, are determined in accordance to Equation (3). This algorithmic procedure permits the evaluation of the mean concentration "seen" by a sensor at any arbitrary receptor space–time point $(\mathbf{x}_r, t_r)$ in accordance to Equation (2). Note that evaluation of the inner product $\overline{C}(\mathbf{x}_r, t_r) = \langle C, h \rangle (\mathbf{x}_r, t_r)$ involves two basic functions: the source distribution $S(\mathbf{x}', t')$ and the detector space–time filtering function $h(\mathbf{x}, t | \mathbf{x}_r, t_r)$. Unfortunately, the source-oriented approach described here is too computationally expensive for use in a Bayesian inferential approach for multiple source reconstruction, because sampling from the posterior distribution for the source parameters (described later in this report) requires a large number of forward source–receptor relationships to be determined, each of which involves the solution of the stochastic differential equation given by Equation (3). Therefore, for the Bayesian inversion of concentration data to be practical, fast and efficient techniques are required for the determination of the source–receptor relationship within the context of sampling in the hypothesis space of the source required for posterior inference.

In view of this, it is much more attractive to apply a receptor-oriented approach for representation of the source–receptor relationship, owing to the fact that this approach provides a fast technique for predicting mean concentration data at a given receptor location for arbitrary hypotheses about the source distribution. As a consequence, we adopt and develop this formulation of the source–receptor relationship, as it provides a fast and reliable framework for Bayesian inference in the context of source reconstruction. To this purpose, we consider a representation for the source–receptor relationship that is dual to the one given by Equation (2) as follows:

$$\begin{aligned} \overline{C}\left(\mathbf{x}_r, t_r\right) &\equiv \int_0^T dt \int_{\mathcal{D}} d\mathbf{x} C\left(\mathbf{x}, t\right) h(\mathbf{x}, t | \mathbf{x}_r, t_r) \\ &= \int_{-\infty}^{t_r} dt' \int_{\mathcal{D}} d\mathbf{x}' C^*\left(\mathbf{x}', t' | \mathbf{x}_r, t_r\right) S(\mathbf{x}', t') \equiv \langle C^*, S \rangle (\mathbf{x}_r, t_r), \end{aligned} \tag{6}$$

where $C^*(\mathbf{x}', t' | \mathbf{x}_r, t_r)$ is an adjunct (dual) concentration at space–time point $(\mathbf{x}', t')$ associated with the sensor concentration data at location $\mathbf{x}_r$ and time t_r (with $t' < t_r$). We note that $C^*(\mathbf{x}', t' | \mathbf{x}_r, t_r)$ is explictly constructed so that it verifies the duality relationship implied by Eq. (6); namely, $\langle C, h \rangle = \langle C^*, S \rangle$.

A backward-time Lagrangian trajectory simulation model, that is dual to the forward-time Lagrangian trajectory simulation model given by Equation (1), can be constructed for the computation of C^* so that it exactly satisfies the duality relationship $\langle C, h\rangle = \langle C^*, S\rangle$. To this end, suppose the backward-time Lagrangian trajectory model is defined as the solution to the following stochastic differential equation:

$$\begin{aligned} d\mathbf{X}^b(t') &= \mathbf{U}^b(t')\,dt', \\ d\mathbf{U}^b(t') &= \mathbf{a}^b(\mathbf{X}^b(t'), \mathbf{U}^b(t'), t')\,dt' + \big(C_0\epsilon(\mathbf{X}^b(t'), t')\big)^{1/2}\,d\mathbf{W}(t'), \end{aligned} \tag{7}$$

where $t' < t_r$ and at any given time t', $(\mathbf{X}^b, \mathbf{U}^b)$ is a point in the phase space along the backward trajectory of the "marked" fluid element (here assumed to be marked or tagged as a fluid particle which at time t_r passed through the spatial volume of the detector at location $\mathbf{x}_r$). The displacement statistics of "marked" fluid elements released from the receptor location $\mathbf{x}_r$ at time t_r can be used to compute $C^*(\mathbf{x}', t'|\mathbf{x}_r, t_r)$ (which is interpreted here as a function of $\mathbf{x}'$ and t'). It can be shown (Thomson [18], Flesch et al. [4]) that C^* obtained from Equation (7) for a detector with the filtering function h and C obtained from Equation (2) for a release from the source density S is exactly consistent with the duality relationship $\langle C, h\rangle = \langle C^*, S\rangle$ if and only if $\mathbf{a}^b$ in Equation (7) is related to $\mathbf{a}$ in Equation (3) through the following gauge transformation:

$$a_i^b(\mathbf{x}, \mathbf{u}, t) = a_i(\mathbf{x}, \mathbf{u}, t) - C_0\epsilon(\mathbf{x}, t)\frac{\partial}{\partial u_i}\ln P_E(\mathbf{u}). \tag{8}$$

This form of the drift coefficient for the backward-time Lagrangian stochastic model is a consequence of Thomson's well-mixed criterion and guarantees that the duality relationship between C^* and C, for any source distribution S and detector response function h, is exactly satisfied. Substituting Equations (4) and (5a) in Equation (8), an explicit form for a_i^b can be obtained as

$$a_i^b = \frac{1}{2}\big(C_0\epsilon\big)\Gamma_{ik}^{-1}\big(u_k - \overline{U}_k\big) + \frac{\phi_i}{P_E}. \tag{9}$$

Note that a_i^b differs from a_i by virtue of a sign change on the first term (or, "fading" memory term) on the right-hand-side of Equation (9). The second term ("drift" term due to spatial inhomogeneity and/or temporal non-stationarity in the velocity statistics of the turbulent flow) on the right remains invariant under a time reversal operation. This suggests that the "fading" memory and "drift" terms transform differently under a time reversal operation.

Model for concentration observations

The models described above provide predictions for the "ideal" mean concentration seen by a sensor at the receptor space–time point $(\mathbf{x}_r, t_r)$. The actual concentration data measured by the sensor will not usually agree with the concentration predicted by the model owing to the noise process imposed on the concentration data, which by its very nature is expected to have a very complicated structure. To this purpose, it is assumed that the actual concentration data available from the sensor array were measured at a finite number of sensor locations and at a finite number of time points at each sensor location. The actual concentration datum $d_{i,j^{(i)}}$ acquired by the sensor at receptor location $\mathbf{x}_{r_i}$ and at time $t_j^{(i)}$ $(i = 1, 2, \ldots, N_d$ and $j = 1, 2, \ldots, N_t^{(i)}$, where N_d is the number of sensors and $N_t^{(i)}$ is the number of time samples

measured at the i-th sensor) is assumed to be the sum of a modeled mean concentration signal $\overline{C}(\mathbf{x}_{r_i}, t_j^{(i)}; \Theta)$ and "noise" $e_{i,j^{(i)}}$, so

$$d_{i,j^{(i)}} = \overline{C}(\mathbf{x}_{r_i}, t_j^{(i)}; \Theta) + e_{i,j^{(i)}}, \tag{10}$$

where Θ is an appropriate parameter vector describing the source distribution S; and, $\overline{C}(\mathbf{x}_r, t_r; \Theta)$ is the modeled mean concentration at location $\mathbf{x}_r$ and time t_r, determined in accordance to Equation (6) for a source distribution characterized by parameter vector Θ. For simplicity of notation, the variables in Equation (10) which are indexed or labelled by $(i, j^{(i)})$ will be ordered in some regular and convenient manner (e.g., lexicographic ordering) and this collection will be indexed by J ($J = 1, 2, \ldots, N$, with $N \equiv \sum_{i=1}^{N_d} N_t^{(i)}$ being the total number of measured concentration data). Then, we can write the observational model as follows:

$$d_J = \overline{C}_J(\Theta) + e_J, \qquad J = 1, 2, \ldots, N, \tag{11}$$

where $\overline{C}_J(\Theta) \equiv \overline{C}(\mathbf{x}_{r_i}, t_j^{(i)}; \Theta)$.

In Equation (11), the e_J is a noise term representing the uncertainty in d_J. In general, e_J consist of errors (e.g, input, stochastic, and measurement) and any real signal in the data that cannot be explained by the model. The random error e_J can be split into four terms as discussed by Rao [19], so

$$e_J = \eta_J^{(1)} + \eta_J^{(2)} + \eta_J^{(3)} + \eta_J^{(4)}. \tag{12}$$

The first term $\eta_J^{(1)}$ of the error corresponds to model error arising from uncertainties in the representation of various physical processes in the dispersion model used to predict the mean concentration. The second term $\eta_J^{(2)}$ describes the input error arising from uncertainties in the values of empirical parameters and/or specification of the input meteorology (initial and boundary conditions) used by the dispersion model. The third term $\eta_J^{(3)}$ of the error is the stochastic uncertainty arising from the turbulent nature of the atmosphere, which gives rise naturally to random concentration fluctuations in hazardous gas releases. Finally, $\eta_J^{(4)}$ describes the noise inherent in the sensor (essentially measurement or instrument error).

Rao [19] discusses the nature of these four types of error with respect to characterization of uncertainties in atmospheric dispersion models, and provides a comprehensive review of sensitivity and/or uncertainty analysis methods that have been used to quantify and reduce them. In this paper, all the various error contributions to the noise term are simply lumped together and denoted by e_J [see Equations (11) and (12)]. It is assumed that the observer does not have a detailed knowledge of the probability distribution of the noise (aggregate error), other than that the observer has an estimate for the expected scale of variation of the noise. More specifically, it is assumed that the noise scale parameter associated with e_J (aggregate error of the J-th concentration observation) is provided in the form of a (finite) variance σ_J^2.

In this paper, it is considered that the model concentration $\overline{C}_J(\Theta)$ in Equation (11) results from the release of a contaminant (e.g., CBR material) from N_s transient point sources. It is assumed that the k-th source was located at $\mathbf{x}_{s,k}$ and that this source was activated (turned on) at time T_b^k and deactivated (turned off) at a later time T_e^k. Furthermore, over the time

period from T_b^k to T_e^k, the source was assumed to be releasing contaminant at a constant emission rate Q_k ($k = 1, 2, \ldots N_s$). Consequently, the (unknown) source distribution is assumed *a priori* to have the following form:

$$S(\mathbf{x}, t) = \sum_{k=1}^{N_s} Q_k \delta\big(\mathbf{x} - \mathbf{x}_{s,k}\big) \Big[U(t - T_b^k) - U(t - T_e^k)\Big], \tag{13}$$

where $\delta(\cdot)$ and $U(\cdot)$ are the Dirac delta and Heaviside unit step functions, respectively. It is convenient to define a source parameter vector θ_{N_s} corresponding to the source distribution S of Equation (13) as $\theta_{N_s} \equiv (\mathbf{x}_{s,1}, T_b^1, T_e^1, Q_1, \ldots, \mathbf{x}_{s,N_s}, T_b^{N_s}, T_e^{N_s}, Q_{N_s}) \in I\!R^{6N_s}$. Furthermore, let $\Theta \equiv (N_s, \theta_{N_s})$. If we substitute Equation (13) into Equation (6), the model concentration $\overline{C}_J(\Theta)$ "seen" by the sensor at location $\mathbf{x}_{rJ}$ and time t_{rJ} is given explicitly by

$$\overline{C}_J(\Theta) = \sum_{k=1}^{N_s} Q_k \int_{T_b^k}^{\min(t_{rJ}, T_e^k)} C^*(\mathbf{x}_{s,k}, t_s | \mathbf{x}_{rJ}, t_{rJ})\, dt_s. \tag{14}$$

The problem of source reconstruction reduces to the following: given the observed vector of concentration data $\mathbf{D} \equiv (d_1, d_2, \ldots, d_N)$, the objective is to estimate N_s and θ_{N_s} or, equivalently, Θ.

Bayesian analysis

To determine the number of sources and reconstruct the characteristics (e.g., location, emission rate) of each source (parameters encapsulated in Θ), given the vector of concentration observations $\mathbf{D}$ obtained from the array of sensors, the techniques of Bayesian inference are employed. The Bayesian framework is attractive in the current context because it provides a rigorous mathematical foundation for making inferences about the source parameters and, as a consequence, provides a rigorous basis for quantifying the uncertainties in the estimated source parameters. Bayesian inference can be obtained from the product rule of probability calculus, the latter of which can be derived rigorously starting with the formulation of a small number of desiderata required to define a rational theory of inference as first enunciated by Cox [20], with a more complete treatment given by Jaynes [21] in his definitive treatise. This formulation leads to the ordinary rules (sum and product rules) of probability calculus and implies that every allowed (consistent) theory for inference must be mathematically equivalent to probability theory, or else inconsistent (no other calculus is admissible for inference that is consistent with the above mentioned desiderata).

The basic relationship quantifying parameter inference is Bayes' rule, which in the context of the current problem can be expressed as follows:

$$p(\Theta|\mathbf{D}, I)p(\mathbf{D}|I) = p(\mathbf{D}|\Theta, I)p(\Theta|I). \tag{15}$$

The factors in Equation (15) are: (1) $p(\Theta|I)$ is the prior probability density function (PDF) of the source parameter vector Θ that encapsulates our state of knowledge of the parameters before the receipt of the concentration measurements; (2) $p(\mathbf{D}|\Theta, I)$ is termed the

likelihood function when considered as a function of Θ, but is known as the sampling distribution when considered as a function of $\mathbf{D}$ (the latter being the probability of observing the concentration data $\mathbf{D}$ when given the source parameters Θ); (3) $p(\mathbf{D}|I)$ is termed the evidence (also, frequently referred to as the prior predictive or the marginal likelihood) ; and, (4) $p(\Theta|\mathbf{D}, I)$ is the posterior probability density function of the parameters Θ of interest, that corresponds to the update of $p(\Theta|I)$ incorporating the knowledge gained about Θ after receipt of the concentration observations $\mathbf{D}$. All the terms here are to be interpreted given the background (contextual) information I (e.g., background meteorology, source–receptor relationship, etc.) that defines the source reconstruction problem. This is the meaning of I after the vertical bar "|" which is used here to denote "conditional upon".

Using Bayesian inference, all information in the measured concentration data relevant to the problem of estimating the source parameter vector Θ is summarized in the posterior PDF $p(\Theta|\mathbf{D}, I)$, which by virtue of Equation (15), is given by

$$p(\Theta|\mathbf{D}, I) \propto p(\mathbf{D}|\Theta, I)p(\Theta|I). \tag{16}$$

It is noted that for parameter estimation, the evidence $p(\mathbf{D}|I)$ is Θ-independent and simply plays the role of a normalization factor; hence, $p(\Theta|\mathbf{D}, I)$ is expressed simply as a proportionality in Equation (16). To proceed further in the specification of the posterior PDF, we now need to assign functional forms for the prior PDF $p(\Theta|I)$ and the likelihood function $p(\mathbf{D}|\Theta, I)$.

First, let us consider the assignment of the prior PDF $p(\Theta|I) \equiv p((N_s, \theta_{N_s})|I)$. To this purpose, we assume the logical independence of the various source parameters; that is to say, knowing the value of the number of sources tells us nothing about the various source characteristics (e.g., location, emission rate), knowing the characteristics of one source tells us nothing about those of another source, knowing the location of a particular source tells us nothing about the emission rate or when the source was turned on/off, etc. In consequence, the prior probability $p((N_s, \theta_{N_s})|I)$ factorizes (by repeated application of the product rule of probability calculus) as follows:

$$p((N_s, \theta_{N_s})|I) = p(N_s|I) \prod_{k=1}^{N_s} p(Q_k|I)p(\mathbf{x}_{s,k}|I)p(T_b^k|I)p(T_e^k|T_b^k, I). \tag{17}$$

We now discuss the assignment of each of the (component) prior distributions in Equation (17). The prior distribution on N_s (number of sources) is chosen to be a truncated Poisson distribution with parameter Γ, where Γ is the expected number of sources, so

$$p(N_s|I) \propto \Gamma^{N_s} \frac{\exp(-\Gamma)}{N_s!} \mathbb{I}_{\{1,2,\ldots,N_{s,\max}\}}(N_s), \tag{18}$$

where $N_{s,\max}$ is the maximum number of sources and $\mathbb{I}_A(x)$ is the indicator function for set A (viz., $\mathbb{I}_A(x) = 1$ if $x \in A$, and $\mathbb{I}_A(x) = 0$ if $x \notin A$). Here and elsewhere, irrelevant normalization constants in functional forms of probability distributions will be ignored. Given N_s, the prior distributions for Q_k, $\mathbf{x}_{s,k}$, T_b^k and T_e^k $(k = 1, 2, \ldots, N_s)$ are each chosen to be uniform over an appropriate domain of definition. More specifically, the prior distribution for Q_k is uniform:

$$p(Q_k|I) = \mathbb{I}_{(0,Q_{\max})}(Q_k)/Q_{\max}, \qquad k = 1, 2, \ldots, N_s, \tag{19}$$

where $Q_{\max}$ is an upper bound for the expected emission rate. The prior distribution for $\mathbf{x}_{s,k}$ is chosen to be uniform over some large region $\mathcal{D} \subset I\!R^3$:

$$p(\mathbf{x}_{s,k}|I) = I\!I_{\mathcal{D}}(\mathbf{x}_{s,k})/V(\mathcal{D}), \qquad k = 1, 2, \ldots, N_s, \tag{20}$$

where $V(\mathcal{D})$ is the volume of the region $\mathcal{D}$. Finally, the prior distributions for T_b^k and T_e^k are assigned uniform distributions with the following forms:

$$p(T_b^k|I) = I\!I_{(0,T_{\max})}(T_b^k)/T_{\max}, \qquad k = 1, 2, \ldots, N_s; \tag{21}$$

and

$$p(T_e^k|T_b^k, I) \propto I\!I_{(T_b^k,T_{\max})}(T_e^k)/\left(T_{\max} - T_b^k\right), \qquad k = 1, 2, \ldots, N_s. \tag{22}$$

Here, $T_{\max}$ is the upper bound on the time at which the source was turned on and turned off. Note that the time that the source is turned off must necessarily occur after it has been turned on, and this information is encoded in form of the prior distribution for T_e^k given by Equation (22), where the distribution is seen to be conditioned on T_b^k. In view of Equations (18) to (22), the prior distribution for Θ given by Equation (17) now assumes the following explicit form (up to a normalizing constant that does not depend on the source parameters):

$$\begin{aligned} p((N_s, \theta_{N_s})|I) \propto \Gamma^{N_s} \frac{\exp(-\Gamma)}{N_s!} I\!I_{\{1,2,\ldots,N_{s,\max}\}}(N_s) \\ \times \prod_{k=1}^{N_s} I\!I_{(0,Q_{\max})}(Q_k) I\!I_{\mathcal{D}}(\mathbf{x}_{s,k}) I\!I_{(0,T_{\max})}(T_b^k) \frac{I\!I_{(T_b^k,T_{\max})}(T_e^k)}{\left(T_{\max} - T_b^k\right)}. \end{aligned} \tag{23}$$

Next, consider the assignment of a functional form for the likelihood function $p(\mathbf{D}|\Theta, I)$. The likelihood function is equivalent to the direct probability for the concentration data $\mathbf{D}$ given the source parameters Θ. In the absence of a detailed knowledge of the noise distribution e_J [cf. Equation (11)], other than that it has a finite variance σ_J^2, the application of the principle of maximum entropy (Jaynes [21]) informs us that a Gaussian distribution is the most conservative choice for the direct probability of the data $\mathbf{D}$ (or, equivalently, for the noise $\mathbf{e} \equiv (e_1, e_2, \ldots, e_N)$). It is important to note that the entropy of the noise PDF (which is equivalent to the likelihood function) is a measure of the size of the basic support set of the distribution or 'volume' occupied by the sensibly probable noise values. Choosing a distribution for the noise that provides the largest support set permitted by the available information allows the largest range of possible variations in the noise values consistent with that information, implying the most conservative estimates for these values. Assigning a Gaussian distribution for the noise using the maximum entropy principle makes no statement about the true (unknown) sampling distribution of the noise (which has been indicated earlier to have a highly complicated structure for the current problem). Rather, it simply represents a maximally uninformative state of knowledge, a state of knowledge that reflects what the observer knows about the true noise in the data (namely, the mean and variance of the noise, with all other properties of the noise being irrelevant to the inference since these are unknown to the observer). From these considerations, the likelihood function for our problem has the following form [in light of Equation (11)]:

$$p(\mathbf{D}|\Theta, I) = \frac{1}{\prod_{J=1}^{N} \sqrt{2\pi}\sigma_J} \exp\left(-\frac{1}{2}\chi^2(\Theta)\right), \tag{24a}$$

where

$$\chi^2(\Theta) \equiv \sum_{J=1}^{N} \left(\frac{d_J - \overline{C}_J(\Theta)}{\sigma_J} \right)^2 . \tag{24b}$$

Inserting Equations (23) and (24) into Equation (16), the posterior distribution for Θ has the following form, where all superfluous constants that are independent of N_s or θ_{N_s} have been absorbed into the constant of proportionality (or, normalization constant):

$$\begin{aligned} p(\Theta|\mathbf{D}, I) \propto & \frac{1}{\prod_{J=1}^{N} \sqrt{2\pi}\sigma_J} \exp\left(-\frac{1}{2} \sum_{J=1}^{N} \left(\frac{d_J - \overline{C}_J(\Theta)}{\sigma_J} \right)^2 \right) \\ & \times \Gamma^{N_s} \frac{\exp(-\Gamma)}{N_s!} \mathbb{I}_{\{1,2,\ldots,N_{s,\max}\}}(N_s) \\ & \times \prod_{k=1}^{N_s} \mathbb{I}_{(0,Q_{\max})}(Q_k) \mathbb{I}_{\mathcal{D}}(\mathbf{x}_{s,k}) \mathbb{I}_{(0,T_{\max})}(T_b^k) \frac{\mathbb{I}_{(T_b^k, T_{\max})}(T_e^k)}{\left(T_{\max} - T_b^k\right)}, \end{aligned} \tag{25}$$

recalling that the source parameter vector $\Theta \equiv (N_s, \theta_{N_s})$. Here, $\overline{C}_J(\Theta)$ is determined in accordance to Equation (14) with the adjunct "concentration" field $C^*(\mathbf{x}', t'|\mathbf{x}_r, t_r)$ predicted using a backward-time Lagrangian stochastic model [cf. Equations (7), (9), and (5b)]. The posterior distribution for Θ provides the full solution for the multiple source reconstruction problem. Inferences on the values of the source parameters are based on upon this posterior distribution. This posterior distribution may be summarized by various statistics of interest such as the posterior mean of each source parameter, say $\theta^i_{N_s}$ (which can be an emission rate Q_k, or source location $\mathbf{x}_{s,k}$, or time at which the source was turned on (off) T_b^k (T_e^k) for the k-th source for a source distribution consisting of N_s sources with $k = 1, 2, \ldots, N_s$):

$$\overline{\theta^i_{N_s}} = \mathcal{E}[\theta^i_{N_s}|\mathbf{D}] \equiv \int \theta^i_{N_s} p((N_s, \theta_{N_s})|\mathbf{D}, I)\, d\theta_{N_s}, \tag{26a}$$

where $\mathcal{E}[\cdot]$ denotes mathematical expectation. A measure of uncertainty of this estimate of $\theta^i_{N_s}$ is the posterior standard deviation

$$\sigma(\theta^i_{N_s}) = \mathcal{E}^{1/2}[(\theta^i_{N_s} - \overline{\theta^i_{N_s}})^2|\mathbf{D}] \equiv \left(\int \left(\theta^i_{N_s} - \overline{\theta^i_{N_s}}\right)^2 p((N_s, \theta_{N_s})|\mathbf{D}, I)\, d\theta_{N_s} \right)^{1/2} . \tag{26b}$$

Alternatively, a p% credible (or, highest posterior distribution) interval that contains the source parameter θ^i with p% probability, with the lower and upper bounds of the interval specified such that the probability density within the interval is everywhere larger than that outside it, can be used as a measure of the uncertainty in the determination of θ^i. Finally, an estimate for the number of sources can be obtained from the maximum a posteriori estimate as follows:

$$\widehat{N}_s = \underset{N_s}{\operatorname{argmax}}\ p(N_s|\mathbf{D}, I), \tag{27}$$

where $p(N_s|\mathbf{D}, I)$ is the posterior probability of the number of sources, given the concentration data and the contextual background information. The posterior probability for the number of sources is a marginal posterior probability, where all the parameters we are not interested in (e.g., θ_{N_s}) are removed, through the process of integrating over these parameters (in a process called marginalization).

Bayesian computation and Markov chains

A perusal of Equation (25) shows that the posterior distribution is highly nonlinear in the source parameters associated with location ($\mathbf{x}_{s,k}$) and with the times the source was turned on (T_b^k) or off (T_e^k), and that explicit evaluations of Equations (26) and (27) (and similar integrals or functionals arising in Bayesian analysis) are impossible. In view of this, we apply posterior sampling for evaluation of these integrals, which is implemented using a Markov chain Monte Carlo (MCMC) algorithm (Wilks et al. [22], Gelman et al. [23]). A MCMC algorithm can be used to generate samples of source distribution models (characterized by Θ) from the posterior distribution in Equation (25). Towards this objective, a Markov chain $\{\Theta^{(t)}\} \equiv \{(N_s^{(t)}, \theta_{N_s^{(t)}}^{(t)})\}$ is constructed whose stationary (or, invariant) distribution is the posterior distribution $p(\Theta|\mathbf{D}, I)$ of the parameters $\Theta = (N_s, \theta_{N_s})$.

All quantities of interest, such as posterior means of various source parameters and various marginal posterior distributions, can be estimated by sample path averages of the Markov process $\{\Theta^{(t)}\}$. For example, the marginal posterior distribution $p(N_s|\mathbf{D}, I)$, which is required in Equation (27) for the inference of N_s, can be estimated as follows:

$$\widehat{p}(N_s|\mathbf{D}, I) = \frac{1}{N_*} \#\big\{t : N_s^{(t)} = N_s\big\} = \frac{1}{N_*} \sum_{t=1}^{N_*} I\!\!I_{\{N_s\}}(N_s^{(t)}), \tag{28}$$

where N_* is the number of samples $\Theta^{(1)}, \Theta^{(2)}, \ldots$ drawn from a sampled realization of the Markov chain and $I\!\!I_{\{N_s\}}(\cdot)$ is used to select those samples $\{\Theta^{(t)}, t = 1, 2, \ldots, N_*\}$ drawn from the Markov chain that correspond to source distribution models with exactly N_s sources (viz., those samples with $N_s^{(t)} = N_s$). Similarly, from the samples generated from such a Markov chain, the posterior means of various source parameters $\theta_{N_s}^i$ [see Equation (26a)] can be estimated by

$$\widehat{\mathcal{E}}[\theta_{N_s}^i|\mathbf{D}] = \frac{\sum_{t=1}^{N_*} \theta_{N_s^{(t)}}^{i(t)} I\!\!I_{\{N_s\}}(N_s^{(t)})}{\sum_{t=1}^{N_*} I\!\!I_{\{N_s\}}(N_s^{(t)})}. \tag{29}$$

The difficulty in the construction of a Markov chain for sampling from the posterior distribution in Equation (25) resides in the fact that N_s is an unknown, requiring the design of a Markov chain that can simultaneously explore both parameter and model spaces. In the current problem, we need to consider a set of models $\mathcal{M} = \{M_{N_s}\}_{N_s=1}^{N_{s,\max}}$ indexed by the integer parameter $N_s \in \mathcal{N}_s \equiv \{1, 2, \ldots, N_{s,\max}\}$, each with a source parameter vector $\theta_{N_s} \in I\!R^{6N_s}$. Here, model M_{N_s} corresponds to all permissible source distributions having exactly N_s sources. The simultaneous exploration of a number of candidate source models involving differing numbers of sources can be realized using the reversible–jump MCMC algorithm that was originally introduced by Green [24] in a Bayesian model determination setting.

The remainder of this section will focus on the design of a reversible–jump MCMC algorithm for multiple source reconstruction, in the case where the number of sources is unknown *a priori*. The MCMC sampling required to address this problem involves three basic components: (1) within-model (or, propagation) moves in the Markov chain involving updating the state of the chain on a hypothesis (source parameter) space of fixed dimension (N_s fixed)

using a standard Metropolis-Hastings update; (2) application of a reversible–jump MCMC for between-model moves in the Markov chain involving "jumps" (creation or annihilation of sources) in the state of the chain between hypothesis (source parameter) spaces of different dimensions corresponding to source distributions having different numbers of sources (N_s variable); and, (3) incorporation of parallel tempering and Metropolis-coupled MCMC to increase the "mixing rate" of states of the Markov chain and to achieve improved computational efficiency. We will now describe each of these components of the Markov chain in the remainder of this section.

Propagation moves and MCMC

The objective of MCMC is to construct a Markov chain whose stationary distribution is one that coincides exactly with the target probability density function that we are trying to sample from [e.g., in our case that target probability distribution is the posterior distribution $p(\Theta|\mathbf{D}, I)$ given in Equation (25)]. In this subsection, we consider the problem of design of a MCMC algorithm that samples from $p(\Theta|\mathbf{D}, I)$ for fixed N_s. The basic MCMC algorithm consists of two components: (1) a transition or proposal distribution function (or transition kernel) $T(\Theta'|\Theta)$; and, (2) an acceptance probability $\alpha(\Theta, \Theta')$. These two components are related as follows: given a chain in the current state $\Theta^{(t)} = \Theta$ at iteration t, a proposed new state $\Theta' = \mathcal{P}(\Theta) = \mathcal{P}\big((N_s, \theta_{N_s})\big) = \big(N_s, \theta'_{N_s}\big)$ is drawn from some proposal (transition) distribution $T(\Theta'|\Theta)$ and this new point is accepted as the new state of the chain at iteration $(t+1)$ with the standard Metropolis-Hastings (M-H) acceptance probability (Gelman et al. [23]) given by

$$\alpha(\Theta, \Theta') = \min\left\{1, \frac{p(\Theta'|\mathbf{D}, I)T(\Theta'|\Theta)}{p(\Theta|\mathbf{D}, I)T(\Theta|\Theta')}\right\}. \tag{30}$$

Note that $\mathcal{P}(\cdot)$ is a pure *propagation* operation (move) that "translates" $\theta_{N_s} \in I\!R^{6N_s}$ to $\theta'_{N_s} \in I\!R^{6N_s}$, while keeping N_s fixed (update move in a fixed dimensional hypothesis space). If the proposal for this propagation move is accepted, then the new state at iteration $(t+1)$ is $\Theta^{(t+1)} = \Theta'$; otherwise, $\Theta^{(t+1)} = \Theta$. Finally, it is noted that the M-H algorithm does not require the normalization of $p(\Theta|\mathbf{D}, I)$ be known [cf. Equation (30)].

To proceed further, it is necessary to specify appropriate forms for the proposal density $T(\Theta'|\Theta)$. For the propagation move, N_s is fixed and the update occurs only for θ_{N_s}. To this purpose, it is useful to partition θ_{N_s} into two blocks of parameters as follows: $\theta_{N_s} = (\theta^1, \theta^2)$ where $\theta^1 \equiv (Q_1, Q_2, \ldots, Q_{N_s}) \in I\!R^{N_s}$ and $\theta^2 = (\mathbf{x}_{s,1}, T_b^1, T_e^1, \ldots, \mathbf{x}_{s,N_s}, T_b^{N_s}, T_e^{N_s}) \in I\!R^{5N_s}$. This particular partitioning of the source parameter vector distinguishes parameters that are related linearly (θ^1) and nonlinearly (θ^2) to the model concentration data as specified in Equation (14). With this partitioning of the source parameter vector, a cycle of two different types of "blocked" component updates (Gibbs step and M-H step) involving θ^1 and θ^2 (respectively) is combined in sequence to form a single iteration of the MCMC sampler for generation of the updated state of the Markov chain.

First, let us consider updating the emission rates Q_k of the various sources ($k = 1, 2, \ldots, N_s$) which we have collected together in θ^1. For the first part of the iteration $(t+1)$, we fix the source parameters in $\theta^2 = \theta^{2(t)}$ (e.g., source locations, activation times, deactivation times) to the values obtained in the previous iteration t and focus on the re-sampling of Q_k ($k = 1, 2, \ldots, N_s$). We choose here to sample the emission rates one-at-time using Gibbs sampling. The Gibbs sampler updates Q_k as a direct draw from the univariate

full conditional posterior distribution $p(Q_k|\theta^1_{-k}, \theta^{2(t)}, \mathbf{D}, I)$, where θ^1_{-k} is used to denote a subvector containing all the components of θ^1 with the exception of Q_k. The Gibbs sampler is simply a special case of the M-H sampler with the proposal density $T(\Theta'|\Theta) = p(Q_k|\theta^1_{-k}, \theta^{2(t)}, \mathbf{D}, I)$ [cf. Equation (30)] for the update of the Q_k component of Θ. With this proposal density, it is straightforward to show that the M-H acceptance probability for the Gibbs transition is exactly unity (all draws from the proposal density are accepted). More specifically, the re-sampling of Q_k $(k = 1, 2, \ldots, N_s)$ for iteration $(t + 1)$ proceeds using a systematic sweep Gibbs sampling strategy as follows:

$$\begin{aligned}
&\text{simulate } Q_1^{(t+1)} \sim p(Q_1|Q_2^{(t)}, Q_3^{(t)}, \ldots, Q_{N_s}^{(t)}, \theta^{2(t)}, \mathbf{D}, I);\\
&\text{simulate } Q_2^{(t+1)} \sim p(Q_2|Q_1^{(t+1)}, Q_3^{(t)}, \ldots, Q_{N_s}^{(t)}, \theta^{2(t)}, \mathbf{D}, I);\\
&\qquad\qquad \vdots\\
&\text{simulate } Q_{N_s}^{(t+1)} \sim p(Q_{N_s}|Q_1^{(t+1)}, Q_2^{(t+1)}, \ldots, Q_{N_s-1}^{(t+1)}, \theta^{2(t)}, \mathbf{D}, I);
\end{aligned} \tag{31}$$

which yields $\theta^{1(t+1)} = (Q_1^{(t+1)}, Q_2^{(t+1)}, \ldots, Q_{N_s}^{(t+1)})$ after N_s cycles. It should be noted that Gibbs sampling can be applied to re-sample Q_k because the univariate conditional distribution $p(Q_k|\theta^1_{-k}, \theta^{2(t)}, \mathbf{D}, I)$ has a particularly simple form from which it is possible to obtain Q_k as a direct draw. Indeed, a perusal of Equations (25) and (14) shows that $p(Q_k|\theta^1_{-k}, \theta^{2(t)}, \mathbf{D}, I)$ is simply a truncated Gaussian distribution with the following form:

$$\begin{aligned}
p(Q_k|\theta^1_{-k}, \theta^{2(t)}, \mathbf{D}, I) \propto \exp\Bigg(&- Q_k\Bigg(\sum_{J=1}^{N}\Bigg[\sum_{\substack{r=1\\ r\neq k}}^{N_s} Q_r f_r^J(\theta^2) - d_J\Bigg] f_k^J(\theta^2)\Bigg)\\
&- \frac{1}{2}Q_k^2\sum_{J=1}^{N} f_k^J(\theta^2)\Bigg) \mathbb{I}_{(0,Q_{\max})}(Q_k),
\end{aligned} \tag{32a}$$

where for simplicity of notation we have defined

$$f_k^J(\theta^2) \equiv \int_{T_b^k}^{\min(t_{rJ}, T_e^k)} C^*(\mathbf{x}_{s,k}, t_s|\mathbf{x}_{rJ}, t_{rJ})\, dt_s. \tag{32b}$$

Following the update of Q_k for iteration $(t+1)$ in accordance to Equation (31), the second part of this iteration involves updating the nonlinear source parameters collected in θ^2 from $\theta^{2(t)}$ to $\theta^{2(t+1)}$. To accomplish this part of the iteration, the emission rate parameters are assumed to be fixed at $\theta^1 = \theta^{1(t+1)}$. For the updating of these parameters, it is no longer possible to use a Gibbs sampler because the resulting univariate full conditional distribution of these parameters cannot be sampled from directly. In view of this, we need to perform a M-H step to update these parameters with an appropriately selected proposal distribution. Towards this objective, we sample each of the nonlinear parameters (locations, activation times, and deactivation times of the various sources) using a M-H step with proposal distribution $T(\theta_k^{2'}|\theta_k^2)$ given as follows:

$$\theta_k^{2'}|\theta_k^2 \sim N(\theta_k^2, \beta_k^2) = \frac{1}{\sqrt{2\pi\beta_k^2}} \exp\left(-\frac{[\theta_k^{2'} - \theta_k^2]^2}{2\beta_k^2}\right), \quad k = 1, 2, \ldots, 4N_s, \tag{33}$$

for valid $\theta_k^{2'}$ (e.g., if the prior information encoded in the prior distribution requires the parameter to lie in a certain interval, then a proposal for this parameter that lies outside this interval is rejected). In Equation (33), θ_k^2 ($k = 1, 2, \ldots, 4N_s$) denotes the k-th component of θ^2. The proposal distribution here provides a candidate $\theta_k^{2'}$ that is a perturbation of the current value of θ_k^2, which is obtained by drawing from a Gaussian distribution with variance β_k^2. Each β_k determines the scale of the proposal steps for the nonlinear parameter θ_k^2, and this scale needs to be carefully selected in order to ensure an appropriate acceptance probability for the proposed (propagation) move. In particular, too small a value for β_k will lead to small propagation steps that are almost all accepted (but result in a very slow exploration of the hypothesis space), whereas too large a value for β_k will result in an excessively high rejection rate for the proposed move (and, in so doing, also severely restrict movement in the hypothesis space). It is a time-consuming process to find an appropriate scale for the proposal distribution for each parameter. In an attempt to circumvent this problem, we use proposal distributions that are mixtures of four Gaussian distributions centered on θ_k^2, each with the form given by Equation (33). However, each Gaussian distribution of this mixture has a different standard deviation β, with the standard deviations chosen in such a manner as to cover several different orders of magnitude. This strategy ensures that a near optimal scale for the step size is used at least some of the time. Because the proposal distribution in Equation (33) is symmetric (for transition between θ_k^2 and $\theta_k^{2'}$), the acceptance probability in Equation (30) for the transition $\theta_k^2 \rightarrow \theta_k^{2'}$, with the posterior probability defined in Equation (25), is given simply by

$$\alpha(\Theta, \Theta') = \min\left\{1, \exp\left(-\frac{1}{2}\Big(\chi^2(\Theta') - \chi^2(\Theta)\Big)\right)\right\}, \tag{34}$$

where Θ' here involves only the change of the k-th component in the subvector θ^2.

Equations (31) and (33) completely specify one iteration involving the sampling of a new candidate state θ_{N_s} for fixed N_s (propagation move). In order to ensure that the Markov chain is reversible, the Gibbs and M-H samplers of Equations (31) and (33) are used systematically to update each parameter in θ^1 and θ^2, respectively, in the forward direction for even iterations and in the reverse direction for odd iterations. In other words, for odd iterations, the procedure for updating θ^1 in Equation (31) is applied backwards in the order $k = N_s, N_s - 1, \ldots, 1$ (and, similarly, for the update of θ^2 for odd iterations).

The propagation move described here only involved the update of θ_{N_s} for a fixed number of sources N_s. In order to move between configurations involving different numbers of sources, we need to use a reversible–jump MCMC procedure that allows updates between states of different dimensions in the hypothesis space (e.g., such as those associated with the creation of a new source or the annihilation of an existing source).

Trans-dimensional jumps and reversible–jump MCMC

A reversible–jump MCMC sampling algorithm (which we use to accomodate between-model moves such as, for example, a change in the number of sources) was first introduced by Green [24] in the context of the development of a methodology for addressing the model selection problem. The reversible–jump MCMC algorithm is very appealing in that it can be considered to be a natural generalization of the standard MCMC algorithm, with the generalization allowing not only moves in a parameter space of fixed dimension but

also "jumps" between model spaces M_{N_s} of different dimensions (viz., involving different numbers of sources N_s). More specifically, suppose a dimension-changing (or, between-model) move of type m is proposed, and the new state Θ' is generated by a deterministic invertible function $g(\Theta, \mathbf{v})$, where $\mathbf{v}$ is a random vector with distribution $f(\mathbf{v})$ [so, $\mathbf{v} \sim f(\mathbf{v})$]. In other words, g is an operator that maps state Θ and the random vector $\mathbf{v}$ into the new state Θ'. Green [24] demonstrated that the acceptance probability of this proposed between-model move m has the following form:

$$\alpha(\Theta, \Theta') = \min\left\{1, \frac{p(\Theta'|\mathbf{D}, I) r_{m'}(\Theta')}{p(\Theta|\mathbf{D}, I) r_m(\Theta) f(\mathbf{v})} \left|\frac{\partial g(\Theta, \mathbf{v})}{\partial(\Theta, \mathbf{v})}\right|\right\}, \tag{35}$$

where m' denotes the reverse move to m and $r_m(\Theta)$ denotes the probability of choosing a move of type m in the state Θ. The final term in the ratio of Equation (35) is the Jacobian which results from the change in variables associated with the dimension-changing move.

For our current problem, we consider two types of dimension-changing moves: namely, a creation move that results in the addition of a single new source to the current source distribution, and an annihilation move that results in the removal of a single existing source from the current source distribution. To be more specific, a creation operator $\mathcal{C}$ generates a between-model move from $(N_s, \theta_{N_s}) \in I\!R^{1+6N_s}$ in model M_{N_s} to $(N_s{+}1, \theta_{N_s+1}) \in I\!R^{1+6N_s+6}$ in model M_{N_s+1}, so $\Theta' = (N_s + 1, \theta_{N_s+1}) \equiv \mathcal{C}(\Theta) = \mathcal{C}\big((N_s, \theta_{N_s})\big)$. Similarly, the reverse move, associated with the annihilation operator $\mathcal{C}^\dagger$, results in an existing source being deleted: so, $\Theta' = (N_s - 1, \theta_{N_s-1}) \equiv \mathcal{C}^\dagger(\Theta) = \mathcal{C}^\dagger\big((N_s, \theta_{N_s})\big)$ where $(N_s - 1, \theta_{N_s-1}) \in I\!R^{1+6N_s-6}$.

To move from model M_{N_s} to M_{N_s+1} using the creation operator $\mathcal{C}$, we propose the generation of a new source at a location $\mathbf{x}_{s,N_s+1}$ with source strength Q_{N_s+1}, emitting material between the activation and deactivation times $T_b^{N_s+1}$ and $T_e^{N_s+1}$, respectively. The "coordinates" of the new source are generated by drawing random samples from some proposal density which we choose to be the prior density for each coordinate: so, from Equations (19) to (22) we sample $\mathbf{x}_{s,N_s+1} \sim p(\mathbf{x}_{s,N_s+1}|I)$, $Q_{N_s+1} \sim p(Q_{N_s+1}|I)$, $T_b^{N_s+1} \sim p(T_b^{N_s+1}|I)$, and $T_e^{N_s+1} \sim p(T_e^{N_s+1}|T_b^{N_s+1}, I)$. Now assemble the parameters describing the new source as $\psi = (\mathbf{x}_{s,N_s+1}, T_b^{N_s+1}, T_e^{N_s+1}, Q_{N_s+1})$, so $\theta_{N_s+1} = \mathcal{C}\big(\theta_{N_s}\big) = \big(\theta_{N_s}, \psi\big)$ for random vector $\psi \in I\!R^6$ whose components are sampled as described above.

Suppose that at any iteration, we propose a creation move with probability $p_{\mathcal{C}}$, with the reverse (annihilation) move having the probability $p_{\mathcal{C}^\dagger}$. The proposal probability for the creation move is then

$$\begin{aligned} q_{\mathcal{C}}(\Theta, \Theta') &= p_{\mathcal{C}} \times p(\mathbf{x}_{s,N_s+1}|I) p(T_b^{N_s+1}|I) p(T_e^{N_s+1}|T_b^{N_s+1}, I) p(Q_{N_s+1}|I) \\ &= p_{\mathcal{C}} \times \frac{1}{V(\mathcal{D})} \frac{1}{T_{\max}} \frac{1}{\big(T_{\max} - T_b^{N_s+1}\big)} \frac{1}{Q_{\max}}. \end{aligned} \tag{36}$$

The probability $q_{\mathcal{C}^\dagger}(\Theta', \Theta)$ for the reverse (annihilation) move is equal to the probability of choosing this move ($p_{\mathcal{C}^\dagger}$) times the probability of picking a particular source from the $N_s + 1$ available sources for annihilation, so

$$q_{\mathcal{C}^\dagger}(\Theta', \Theta) = p_{\mathcal{C}^\dagger} \frac{1}{N_s + 1}. \tag{37}$$

In view of Equation (25), the ratio of posterior distributions in Equation (35) for $\Theta' = \mathcal{C}(\Theta)$ has the following form:

$$\frac{p(\Theta'|\mathbf{D}, I)}{p(\Theta|\mathbf{D}, I)} = \exp\left(-\frac{1}{2}\Big(\chi^2(\Theta') - \chi^2(\Theta)\Big)\right) \times \frac{p(N_s+1|I)}{p(N_s|I)} \frac{1}{V(\mathcal{D})} \frac{1}{T_{\max}} \frac{1}{\big(T_{\max} - T_b^{N_s+1}\big)} \frac{1}{Q_{\max}}. \tag{38}$$

Finally, if we substitute Equations (36), (37), and (38) in Equation (35), and note that for the creation operator $\mathcal{C}$ the Jacobian term in Equation (35) is simply unity, the acceptance probability $\alpha(\Theta, \Theta')$ for the creation move $\Theta' = \mathcal{C}(\Theta)$ is given by

$$\alpha(\Theta, \Theta') = \min\left\{1, \frac{p(N_s+1|I)}{p(N_s|I)} \frac{p_{\mathcal{C}^\dagger}}{p_{\mathcal{C}}} \frac{1}{N_s+1} \exp\left(-\frac{1}{2}\Big(\chi^2(\Theta') - \chi^2(\Theta)\Big)\right)\right\}. \tag{39}$$

As demonstrated by Green [24], a sufficient condition to ensure reversibility of the trans-dimensional Markov chain is for the acceptance ratio for the reverse move to be given by the reciprocal of that for the forward move. In the present context, the condition of detailed balance required for Markov chain reversibility implies that the acceptance probability, associated with the annihilation move $\Theta' = \mathcal{C}^\dagger(\Theta)$ for the removal of a source (from a source distribution containing N_s sources), must have the following form:

$$\alpha(\Theta, \Theta') = \min\left\{1, \frac{p(N_s-1|I)}{p(N_s|I)} \frac{p_{\mathcal{C}}}{p_{\mathcal{C}^\dagger}} N_s \exp\left(-\frac{1}{2}\Big(\chi^2(\Theta') - \chi^2(\Theta)\Big)\right)\right\}. \tag{40}$$

It only remains now to specify the probabilities $p_{\mathcal{C}}$ and $p_{\mathcal{C}^\dagger}$ for creation and annihilation moves, respectively. Of course, with probabilities specfied for the creation and annihilation moves, the probability $p_{\mathcal{P}}$ for the remaining propagation move (source parameters updated for a hypothesis space of fixed dimension) is determined as $p_{\mathcal{P}} = 1 - p_{\mathcal{C}} - p_{\mathcal{C}^\dagger}$. Furthermore, it is useful to allow the probabilities $p_{\mathcal{C}}$ and $p_{\mathcal{C}^\dagger}$ to depend on the number of sources N_s in the current state. To indicate explicitly this dependence on N_s we will augment the notation, and express the probability of a creation and annihilation move as $p_{\mathcal{C}}^{N_s}$ and $p_{\mathcal{C}^\dagger}^{N_s}$, respectively. For $N_s = 1$, $p_{\mathcal{C}^\dagger}^{N_s} = 0$ because at least one source must be responsible for the concentration measured by the array of sensors. Also, for $N_s = N_{s,\max}$, $p_{\mathcal{C}}^{N_s} = 0$ for otherwise the preassigned maximum number of sources that may be responsible for the measured concentration will be exceeded. For all other cases, the probabilities for creation and annihilation moves will be specified as follows:

$$p_{\mathcal{C}}^{N_s} = \frac{1}{2}\min\left\{1, \frac{p(N_s+1|I)}{p(N_s|I)}\right\}, \tag{41a}$$

$$p_{\mathcal{C}^\dagger}^{N_s+1} = \frac{1}{2}\min\left\{1, \frac{p(N_s|I)}{p(N_s+1|I)}\right\}. \tag{41b}$$

This particular choice for $p_{\mathcal{C}}^{N_s}$ and $p_{\mathcal{C}^\dagger}^{N_s+1}$ guarantees that $p_{\mathcal{C}}^{N_s} p(N_s|I) = p_{\mathcal{C}^\dagger}^{N_s+1} p(N_s+1|I)$ and ensures that the probability of a dimension-changing move at each iteration is between

0.5 and 1 (Green [24]). Finally, it should be noted that with the dependence of $p_{\mathcal{C}}$ and $p_{\mathcal{C}^\dagger}$ on N_s as in Equation (41), the ratios of these two probabilities in Equations (39) and (40) need to be interpreted as follows: namely, $p_{\mathcal{C}^\dagger}/p_{\mathcal{C}} \rightarrow p_{\mathcal{C}^\dagger}^{N_s+1}/p_{\mathcal{C}}^{N_s}$ in Equation (39) and $p_{\mathcal{C}}/p_{\mathcal{C}^\dagger} \rightarrow p_{\mathcal{C}}^{N_s-1}/p_{\mathcal{C}^\dagger}^{N_s}$ in Equation (40).

Parallel tempering and Metropolis-coupled MCMC

To overcome problems associated with a Markov chain that is slowly mixing in the hypothesis space and which can cause the chain to be stuck in a local minimum, we have implemented a form of parallel tempering to obtain a Metropolis-coupled MCMC algorithm that improves the speed at which the hypothesis space is explored. The Metropolis-coupled MCMC algorithm has been described by Geyer [25] and is based on the idea of using a series of transition densities $T_1(\Theta'|\Theta), T_2(\Theta'|\Theta), \ldots, T_r(\Theta'|\Theta)$ with corresponding (unnormalized) invariant (stationary) distributions $p_1(\Theta|\mathbf{D}, I), p_2(\Theta|\mathbf{D}, I), \ldots, p_r(\Theta|\mathbf{D}, I)$. In our current application of Metropolis-coupled MCMC, we run r Markov chains in parallel to give samples $\Theta_{[i]}^{(t)}$ $(i = 1, 2, \ldots, r)$, where chain i has the stationary distribution given by

$$p_i(\Theta|\mathbf{D}, I) = p(\Theta|\mathbf{D}, I)p^{\lambda_i}(\mathbf{D}|\Theta, I), \quad i = 1, 2, \ldots, r, \tag{42}$$

where $\lambda_i \in [0, 1]$ $(i = 1, 2, \ldots, r)$ is an increasing sequence (viz., $\lambda_i < \lambda_j$ for $i < j$) with $\lambda_1 = 0$ and $\lambda_r = 1$.

The parameter λ_i in Equation (42) is interpreted as a tempering parameter. This parameter is used to raise the likelihood function to the λ_i power to give a modified posterior proportional to $p(\Theta|I)p^{\lambda_i}(\mathbf{D}|\Theta, I)$. It is noted that $\lambda_r = 1$ is associated with the desired target posterior probability distribution [cf. Equation (25)] that we want to sample from. The other simulations correspond to a "ladder" of modified posterior distributions indexed by i. As the tempering parameter increases from zero to one, the effects of the concentration data are introduced slowly through the "softened" likelihood function $p^{\lambda_i}(\mathbf{D}|\Theta, I)$ $[\lambda_i \in [0, 1)$ for $i = 1, 2, \ldots, r - 1]$. In particular, $\lambda_i < 1$ implies that the corresponding modified posterior distribution is broader (or flatter) than the actual posterior distribution, allowing the states to move more freely in the hypothesis space.

Let us denote the state of the i-th Markov chain at iteration t by $\Theta_{[i]}^{(t)}$ $(i = 1, 2, \ldots, r)$. In this form of Metropolis-coupled MCMC using parallel tempering, the r Markov chains are run simultaneously and at each iteration there is a prescribed probability (or, equivalently, on average every N_{iter} iterations) that a pair of adjacent chains (say, i and $i + 1$ with $i = 1, 2, \ldots, r - 1$) is randomly selected, and a proposal is made to swap the states of these two chains. More specifically, if at iteration t, a proposed swapping operation between chains i and $i + 1$ is accepted, then we swap the states of the chains $\Theta_{[i]}^{(t)} \rightarrow \Theta_{[i+1]}^{(t)}$ and $\Theta_{[i+1]}^{(t)} \rightarrow \Theta_{[i]}^{(t)}$ with an acceptance probability for this swap given by

$$\alpha_{\text{swap}} = \min\left\{1, \frac{p_i(\Theta_{[i+1]}^{(t)}|\mathbf{D}, I)p_{i+1}(\Theta_{[i]}^{(t)}|\mathbf{D}, I)}{p_i(\Theta_{[i]}^{(t)}|\mathbf{D}, I)p_{i+1}(\Theta_{[i+1]}^{(t)}|\mathbf{D}, I)}\right\}. \tag{43}$$

This swap enables the exchange of information across the population of r parallel simulations. More specifically, this allows the chain associated with the desired target posterior

distribution (viz., that corresponding to $\lambda_r = 1$) to sample from remote regions of the posterior distribution, which in turns facilitates chain mobility in the hypothesis space and ensures a more reliable relaxation of the chain into the (potentially exponentially tiny) regions of this space where the posterior probability mass is concentrated. We apply the Metropolis-coupled MCMC algorithm described here with $r = 21$. The tempering parameters λ_i used on this ladder of parallel simulations is uniformly spaced between 0 and 1.

Tests with synthetic concentration data

In this section, we present the results of the proposed algorithm for multiple source reconstruction for two cases: namely, case 1 involves two unknown sources, whereas case 2 involves three unknown sourcess. In each of these cases, simulated concentration data will be generated and the Bayesian inference scheme described above will be applied to reconstruct the unknown source parameters. It should be stressed that currently there are no available real concentration data sets that can be used to test our proposed algorithm for an unknown number of multiple sources; hence, the necessity to use simulated concentration data sets instead. However, it should be noted that to address this deficiency, Technical Panel 9 (TP-9) of The Technical Cooperation Program (TTCP) Chemical, Biological and Radiological Defense (CBD) Group has proposed the design and implementation of a cooperative **FU**sing **S**ensor **I**nformation from **O**bserving **N**etworks (FUSION) Field Trial 2007 (FFT 07). This field experiment will be conducted at Tower Grid on U.S. Army Dugway Proving Ground (DPG) in September 2007. The objective of this field experiment will be to provide a comprehensive tracer dispersion and meteorological dataset suitable for testing current and future chemical and biological sensor data fusion algorithms.

In view of this, we will simulate concentration data sets that conform with the proposed design of FFT 07. It should be stressed that the terrain at Tower Grid is flat and smooth with a uniform upwind fetch of more than 20 km. The vegetation at this site consists of either widely spaced low-lying shrub (less than 0.5 m in height) interspersed with bare ground, or very short grass. The aerodynamic roughness length, z_0, for the site is about 0.015 m (Yee et al. [26]). Owing to the horizontal homogeneity of the site, the mean wind flow and turbulence statistics will be assumed to be horizontally homogeneous and stationary. Given the short times and distances over which we will be modeling dispersion, this assumption is quite acceptable (and, indeed, at these distances the dispersion is assumed to occur entirely within the atmospheric surface layer only). To simulate the concentration seen by the array of sensors, we use the forward-time LS model given by Equations (3) and (5) with the choice of the Kolmogorov constant $C_0 = 4.8$, the latter value having been obtained by calibrating the LS model against concentration data measured during a benchmark field experiment Project Prairie Grass (Wilson et al. [27]). The wind flow statistics required by the LS model are prescribed as follows in accordance primarily with well-known surface-layer relations based on Monin-Obukhov theory (Stull [28]).

Let x, y, and z represent coordinates in the alongwind, crosswind, and vertical directions, respectively. The instantaneous (mean) velocities in these three directions will be denoted by u (U), v (V), and w (W), respectively. Furthermore, u', v', and w' are departures of u, v, and w from its mean values U, V, and W, respectively. Then, the stability-dependent

mean wind speed profile in the atmospheric surface layer has the following form:

$$U(z) = \frac{u_*}{k_v}\left[\ln\left(\frac{z}{z_0}\right) - \psi\left(\frac{z}{L}\right)\right], \tag{44}$$

where u_* is the friction velocity, $k_v \approx 0.4$ is von Karman's constant, L is the Obukhov length scale and ψ is the stability function. For unstable stratification ($L < 0$), ψ assumes the form

$$\psi\left(\frac{z}{L}\right) = 2\ln\left(\frac{1+\chi}{2}\right) + \ln\left(\frac{1+\chi^2}{2}\right) - 2\arctan(\chi) + \frac{\pi}{2}, \qquad L < 0, \tag{45a}$$

where

$$\chi \equiv \left(1 - \frac{16z}{L}\right)^{1/4}; \tag{45b}$$

and, for stable stratification ($L > 0$)

$$\psi\left(\frac{z}{L}\right) = \frac{4.7z}{L}. \tag{45c}$$

Note that the mean wind direction is assumed to be along the x-axis, so $V = 0$.

The parametrizations we use for the turbulence velocity flow statistics follow those suggested by Rodean [29]. For the coordinate system used here where the x-axis is chosen to lie along the mean wind direction, the velocity covariances $\overline{u'v'}$ and $\overline{v'w'}$ are zero in the atmospheric surface layer; and, the only remaining velocity covariance $\overline{u'w'}$ is assumed to be constant and equal to $-u_*^2$ (by definition). The alongwind and crosswind velocity variances σ_u^2 and σ_v^2, respectively, are given by

$$\frac{\sigma_u^2}{u_*^2} = \frac{\sigma_v^2}{u_*^2} = 4.5\left(1 - \frac{z}{H}\right)^{3/2}, \qquad L > 0, \tag{46a}$$

and

$$\frac{\sigma_u^2}{u_*^2} = \frac{\sigma_v^2}{u_*^2} = 4.5\left(1 - \frac{z}{H}\right)^{3/2} + 0.6\left(\frac{H}{|L|}\right)^{2/3}, \qquad L < 0. \tag{47b}$$

where H is the boundary-layer height. The vertical velocity variance σ_w^2 is given by

$$\frac{\sigma_w^2}{u_*^2} = 2.0\left(1 - \frac{z}{H}\right)^{3/2}, \qquad L > 0, \tag{48a}$$

and

$$\frac{\sigma_w^2}{u_*^2} = \left[1.6\left(1 - \frac{z}{H}\right)^{3/2} - 3.0\left(\frac{z}{L}\right)\left(1 - 0.98\frac{z}{H}\right)^{3/2}\right]^{2/3}, \qquad L < 0. \tag{48b}$$

Finally, the dissipation ϵ of turbulence kinetic energy is parameterized as follows:

$$\epsilon(z) = \frac{u_*^3}{k_v z}\left(1 + 3.7\frac{z}{L}\right)\left(1 - 0.85\frac{z}{H}\right)^{3/2}, \qquad L > 0, \tag{49a}$$

and

$$\epsilon(z) = \frac{u_*^3}{k_v z}\left(\left(1 + 0.75\frac{z}{|L|}\right)\left(1 - 0.85\frac{z}{H}\right)^{3/2} + 0.3\frac{z}{|L|}\right), \qquad L < 0. \tag{49b}$$

For the two examples used to test the Bayesian inference algorithm for source reconstruction, the concentration measured by a particular sensor in the array will result from the superposition of two or more partially overlapping plumes produced by continuously emitting sources. It is noted that for multiple continuous point sources, $T_b^k \to -\infty$ and $T_e^k \to \infty$ [$k = 1, 2 \ldots, N_s$; cf. Equations (13) and (14)]. In consequence, in this case, the only relevant parameters are the source location $\mathbf{x}_{s,k}$ and the emission rate Q_k for each source ($k = 1, 2, \ldots, N_s$). Furthermore, it is noted that the concentration field C and adjunct concentration field C^* are independent of time [viz., $C(\mathbf{x}, t) = C(\mathbf{x})$ and $C^*(\mathbf{x}, t) = C^*(\mathbf{x})$]. Finally, for the simulations, it is assumed that all the sources are emitting at ground level ($z = 0$) and that this is known *a priori* (viz., this knowledge is considered to be part of the background information I). As a result, the unknown location parameters for each source are its alongwind (x_s) and crosswind (y_s) positions, only.

Example 1: two unknown sources

For the first example, we synthesized artificial concentration data for the case of two sources which were located upwind of an array consisting of 42 detectors arranged as shown in Figure 1. Each detector was placed at a height of 1.5 m above ground level. The two ground level continuously emitting sources were located at $(x_s, y_s) = (-50.0, 0.0)$ m and at $(-250.0, 0.0)$ m and each source had an emission rate of $Q \equiv q_s = 1.0$ g s^{-1}. The flow statistics for the atmospheric surface layer (through which the dispersion occurred) require three surface layer parameters: u_*, L, and z_0 which for the simulations assumed the following values: $u_* = 0.25$ m s^{-1}, $L = 10000$ m, and $z_0 = 0.015$ m (roughness height for Tower Grid at U.S. Army Dugway Proving Ground). It is assumed that the boundary-layer height was $H = 1000$ m. The concentration data generated using the forward-time LS model (with these flow statistics as input) were embedded within white and normally distributed noise with a standard deviation equal to 10% of the true concentration amplitude.

We applied our proposed algorithm for source reconstruction to the simulated concentration data. We perform Metropolis-coupled MCMC sampling with $r = 21$ chains with swaps between these chains attempted at every $N_{\text{iter}} = 25$ iterations on average. The proposed algorithm randomly initializes all unknown parameters (source location, emission rate) in accordance to the prescribed prior distributions for these parameters [cf. Equations (19) and (20)] with $Q_{\max} = 100$ kg s^{-1} and $\mathcal{D} = [-2000, -25]$ m $\times$ $[-500, 500]$ m providing the prior bounds on the emission rate q_s and on the source location (x_s,y_s), respectively. The initial number of sources was randomly assigned in accordance to the prior distribution for N_s given by Equation (18), where $N_{s,\max} = 4$ (maximum allowable number of sources) and the hyperparameter Γ (expected number of sources) is initialized to 1 (hence, the prior distribution favors the wrong choice for the number of sources). For this example, the MCMC algorithm was run for 100,000 iterations, with the first 50,000 iterations corresponding (conservatively) to the burn-in with the result that these samples were discarded from the analysis. The remaining 50,000 post burn-in samples were used for the posterior inference.

Figure 2 (top) displays changes in N_s (number of sources) as a function of the iteration number for the first 500 post burn-in iterations of the reversible–jump MCMC sampler. Note the dimension-changing moves involving creation of a new source (e.g., transitions from $N_s = 2$ to $N_s = 3$ or from $N_s = 3$ to $N_s = 4$) or annihilation of an existing source

(e.g., transitions from $N_s = 4$ to $N_s = 3$ or from $N_s = 3$ to $N_s = 2$). Interestingly, after convergence of the Markov chain to the stationary distribution, moves across models from $N_s = 2$ to $N_s = 1$ (and, vice-versa) do not occur. Indeed, Figure 2 (bottom) which exhibits the probability distribution, $p(N_s) \equiv p(N_s|\mathbf{D}, I)$, for the number of sources estimated from the 50,000 post burn-in samples, shows that the hypothesis $N_s = 1$ source is excluded. The simulations settle in a distribution which favors equally (approximately or better) the hypotheses $N_s = 2$ or 3, with a smaller probability for $N_s = 4$.

Figure 3 exhibits histograms of the alongwind location x_s (top panel), crosswind location y_s (middle panel), and emission rate q_s (bottom panel). These histograms were constructed from the subset of post burn-in samples consisting of two source objects (i.e., samples with $N_s = 2$). The bimodality of the histogram of x_s with mode locations at $x_s \approx -50$ m and $x_s \approx -250$ m indicates that the alongwind positions of the two sources have been correctly identified. Because both sources are located at $y_s = 0$ m and have an emission rate $q_s = 1$ g s^{-1}, the single mode in the histograms at $y_s \approx 0$ m and $q_s \approx 1.0$ g s^{-1} indicates that these source parameters have been correctly inferred for the two sources. The samples corresponding to each mode in the histogram for x_s can be extracted separately. Figures 4 and 5 display the first 12,000 samples obtained from each of these two modes in the form of trace histories of the source parameters (x_s, y_s, q_s) corresponding to each mode. It can be seen that the source parameters associated with each mode (and, corresponding to a particular source object) are estimated well. Finally, it should be stressed that labels used for the source objects in Figures 4 and 5 [e.g., source object 1 is identified to be the source at $(x_s, y_s) = (-250, 0)$ m, whereas source object 2 is the source at $(x_s, y_s) = (-50, 0)$ m] are completely arbitrary here because the posterior probability distribution of the source parameters [cf. Equation (25)] is invariant under a reordering (relabelling) of the identifiers used for each source object. This degeneracy simply corresponds to different (but equivalent) identifications of what is meant by source 1, source 2, etc.

Figures 6 and 7 display histograms of the alongwind location x_s (top panel), crosswind location y_s (middle panel), and emission rate q_s (bottom panel) which were obtained, respectively, from the subset of post burn-in samples consisting of three ($N_s = 3$) and four ($N_s = 4$) source objects. Note from both these figures that there really only exists two modes in the histograms of x_s, and it is seen that these modes correspond to the correct alongwind locations of the two sources. Furthermore, the extra source object in the histograms in Figure 6 and the two extra source objects in the histograms of Figure 7 are randomly distributed in $\mathcal{D}$ (viz., are not concentrated into a significant cluster or clusters of points). Interestingly, these extra source objects have emission rates associated with the mode in the histogram of q_s at ≈ 0 g s^{-1}. The mode in the histogram of q_s at ≈ 1.0 g s^{-1} is associated with the modes in the histogram of x_s at ≈ -250 m and ≈ -50 m and the mode in the histogram of y_s at ≈ 0 m. Hence, even the source reconstruction based on samples with $N_s = 3$ or $N_s = 4$ only give two main clusters of points in $\mathcal{D}$ located at the true positions of the two sources and these two main clusters are associated with the mode in the histogram of q_s at ≈ 1.0 g s^{-1} (corresponding to the true emission rate of the two sources).

From Figures 3, 6, and 7 corresponding to the case $N_s = 2$, 3 and 4, respectively, we see that there exists two clusters in which the samples are concentrated. These two clusters are associated with two source objects. We have obtained all samples for $N_s = 2$, 3, and 4 for each of these two clusters and plotted histograms of the parameters $\{x_s, y_s, q_s\}$ for each

detected source object in Figures 8 and 9. The mean, standard deviation, and lower and upper bounds for the 95% highest posterior density (HPD) interval of the parameters for each detected source object were calculated from the samples for $N_s = 2$, 3, and 4 contained in each of the two identified clusters and the results are summarized in Table 1. From this information, we see that the two source objects have been characterized to a reasonable accuracy.

Example 2: three unknown sources

For our second example, we synthesized artificial concentration data using a forward-time LS model for the case of three sources which were located upwind of an array consisting of 66 detectors arranged as shown in Figure 10. Each detector was placed at a height of 1.5 m above ground level. The three ground level continuously emitting sources were located at $(x_s, y_s) = (-50.0, 0.0)$ m, $(-125.0, -20.0)$ m and $(-300.0, 20.0)$ m and each source had an emission rate of $Q \equiv q_s = 1.0$ g s^{-1}. The properties of the atmospheric surface layer, through which the dispersion from the three sources occurred, was the same as that for the first example above. The synthetic concentration data generated using the forward-time LS model were corrupted with white and normally distributed noise with a standard deviation equal to 10% of the true concentration amplitude.

We applied our proposed algorithm for source reconstruction to the simulated concentration data. The Metropolis-coupled reversible–jump MCMC algorithm used the same parameters as described previously for Example 1, with the following exceptions. Rather than specifying the prior distribution for Q as in Equation (19), we have used instead

$$p(Q_k|I) = (1-\gamma)\delta(Q_k) + \gamma \mathbb{I}_{(0,Q_{\max})}(Q_k)/Q_{\max}, \qquad k = 1, 2, \ldots, N_s, \tag{50}$$

where γ is an intermittency factor, defined as the probability that the source is turned on (viz., $Q_k > 0$). Equation (50), unlike Equation (19), allows for the hypothesis that any given source in the domain can be turned off ($Q_k = 0$) with a finite probability ($= 1 - \gamma$). For the source reconstruction here, we have chosen the hyperparameter $\gamma = 0.2$. The initial number of sources was randomly assigned in accordance to the prior distribution for N_s given by Equation (18), where $N_{s,\max} = 8$ (maximum allowable number of sources) and the hyperparameter Γ (expected number of sources) is initialized to 1. In consequence, our initial specification for the prior distribution for N_s favors the wrong choice for the actual number of sources. We note that the MCMC sampler moved quickly from the low likelihood of the starting point to a region in the hypothesis space with higher likelihood. As in Example 1, the MCMC algorithm was run for 100,000 iterations, with the first 50,000 iterations corresponding to the burn-in with the result that these samples were discarded from the analysis. The remaining 50,000 post burn-in samples were used for the posterior inference, the results of which will be presented below.

An estimate of the posterior distribution for N_s, obtained from a histogram of the number of samples obtained in each subspace of different dimension, is exhibited in Figure 11. We note that the most probable number of source objects is $N_s = 3$, which corresponds to the correct number of sources. This result is obtained in spite of the fact that the prior distribution for N_s was initialized with $\Gamma = 1$ (*a priori* expected number of sources), implying that the algorithm is not sensitive to this hyperparameter. The information embodied in the concentration data was sufficient to move the simulations towards the more complex model

with $N_s = 3$ sources. Interestingly, it was found that with the prior specification for Q_k given by Equation (50), many samples for more complex models involving $N_s > 3$ sources had the additional sources turned off (viz., $Q_k = 0$). It should be noted that for samples involving N_s sources, but with N_d of these sources turned off, the source model here is considered to have $N_s - N_d$ sources, and not N_s (owing to the fact that sources that are turned off do not contribute to the model concentration seen at the detectors). In consequence, some of the samples corresponding to the mode of $p(N_s) \equiv p(N_s|\mathbf{D}, I)$ in Figure 11 at $N_s = 3$ corresponds to more complex models involving more than 3 sources, but with the additional sources turned off.

Figure 12 displays histograms of the source parameters $\{x_s, y_s, q_s\}$ constructed from all samples corresponding to $N_s = 3$ (mode for $p(N_s)$ in Figure 11). Furthermore, we plot the samples obtained for the case $N_s = 3$, projected onto the (x_s, y_s)-subspace in Figure 13. Both Figures 12 and 13 indicate that samples tend to cluster in three main regions of the hypothesis space, which can be identified with three source objects. Histograms of the parameters associated with each of these three source objects are exhibited in Figures 14, 15 and 16. The mean, standard deviation, and lower and upper bounds for the 95% HPD interval of the source parameters for each identified source object are summarized in Table 2. Comparing these values of the parameters of the recovered source objects with the parameters for the three actual sources, we see that the algorithm has adequately recovered the true parameters for each source object, within the stated errors. Not surprisingly, the parameters for the source furtherest away from the array of sensors were determined with the largest uncertainty (as measured either by the standard deviation, or the 95% HPD interval).

Conclusions

In this paper, we have presented a Bayesian inference approach for multiple source reconstruction from a limited number of noisy concentration data obtained from an array of sensors for the case where the number of sources is unknown *a priori*. In this approach, we use a model that relates the source distribution to the concentration data (source–receptor relationship), and then apply Bayesian probability theory to formulate the posterior density function for the source parameters including the number of sources. The evaluation of the posterior density function and of its features of interest requires a numerical procedure. To this end, the computational algorithm required here is implemented using a reversible–jump Markov chain Monte Carlo method to draw samples from the posterior density function. We showed how to design creation and annihilation moves that allow the Markov chain to jump between hypothesis spaces corresponding to different numbers of sources in the source distribution.

The new methodology has been sucessfully applied to simulated concentration data corresponding to two and three source cases. It is shown that the proposed method performs well: the number of sources is correctly identified using the procedure, and the parameters (e.g., location, emission rate) that characterize each identified source are reliably estimated. In addition, the methodology provides a rigorous determination of the uncertainty (e.g., standard deviation, credible intervals) in the inference of the source parameters, hence extending the potential of the methodology as a tool for quantitative source reconstruction.

References

1. Hanna, S. R., Chang, J. and Strimaitis, D. G. (1990). Uncertainties in Source Emission Rate Estimates Using Dispersion Models. *Atmos. Environ.*, 24A, 2971–2980.

2. Gordon, R., Leclerc, M., Schuepp, P. and Brunke, R. (1988). Field Estimates of Ammonia Volatilization From Swine Manure by a Simple Micro-meteorological Technique. *Can. J. Soil Sci.*, 68, 369–380.

3. Wilson, J. D. and Shum, W. K. N. (1992). A Re-examination of the Integrated Horizontal Flux Method for Estimating Volatilization from Circular Plots. *Agric. Forest Meteorol.*, 57, 281–295.

4. Flesch, T. K., Wilson, J. D. and Yee, E. (1995). Backward-time Lagrangian Stochastic Dispersion Models and Their Application to Estimate Gaseous Emissions. *J. Appl. Meteorol.*, 34, 1320–1332.

5. Kaharabata, S. K., Schuepp, P. H. and Desjardins, R. L. (2000). Source Strength Determination of a Tracer Gas Using an Approximate Solution to the Advection-Diffusion Equation for Microplots. *Atmos. Environ.*, 34, 2343–2350.

6. Robertson, L. and Langner, J. (1998). Source Function Estimate by Means of Variational Data Assimilation Applied to the ETEX-I Tracer Experiment. *Atmos. Environ.*, 32, 4219–4225.

7. Nodop, K., Connolly, R. and Girardi, F. (1998). The Field Campaigns of the European Tracer Experiment (ETEX): Overview and Results. *Atmos. Environ.*, 32, 4095–4108.

8. Seibert, P. and Stohl, A. (1999). Inverse Modeling of the ETEX-1 Release With a Lagrangian Particle Model. Third GLOREAM Workshop, Ischia, Italy, 10 pp.

9. Skiba, Y. N. (2003). On a Method of Detecting the Industrial Plants Which Violate Prescribed Emission Rates. *Ecological Modelling*, 159, 125–132.

10. Khapalov, A. V. (1994). Localization of Unknown Sources for Parabolic Systems on the Basis of Available Observations. *Int. J. Systems Sci.*, 25, 1305–1322.

11. Alpay, M. E. and Shor, M. H. (2000). Model-Based Solution Techniques for the Source Localization Problem. *IEEE Trans. Control Sys. Tech.*, 8, 895–904.

12. Matthes, J., Gröll, L. and Keller, H. B. (2005). Source Localization by Spatially Distributed Electronic Noses for Advection and Diffusion. *IEEE Trans Signal Processing*, 53, 1711–1719.

13. Pudykiewicz, J. A. (1998). Application of Adjoint Tracer Transport Equations for Evaluating Source Parameters. *Atmos. Environ.*, 32, 3039–3050.

14. Yee, E. (2005). Probabilistic Inference: An Application to the Inverse Problem of Source Function Estimation. The Technical Cooperation Program (TTCP) Chemical and Biological Defence (CBD) Group Technical Panel 9 (TP-9) Annual Meeting, Defence Science and Technology Organization, Melbourne, Australia.

15. Yee, E. (2006). A Bayesian Approach for Reconstruction of the Characteristics of a Localized Pollutant Source From a Small Number of Concentration Measurements Obtained by Spatially Distributed "Electronic Noses". Russian-Canadian Workshop on Modeling of Atmospheric Dispersion of Weapon Agents, Karpov Institute of Physical Chemistry, Moscow,

Russia.

16. Keats, A., Yee, E. and Lien, F.-S. (2007). Bayesian Inference for Source Determination With Applications to a Complex Urban Environment. *Atmos. Environ.*, 41, 465–479.

17. Yee, E. (2007). Bayesian Probabilistic Approach for Inverse Source Determination From Limited and Noisy Chemical or Biological Sensor Concentration Measurements. SPIE Defence & Security Symposium, Chemical and Biological Sensing VIII, Orlando World Center Marriott Resort and Convention Center, Orlando, Florida, USA, 12 pp.

18. Thomson, D. J. (1987). Criteria for the Selection of Stochastic Models of Particle Trajectories in Turbulent Flows. *J. Fluid Mech.*, 180, 529–556.

19. Rao, K. S. (2005). Uncertainty Analysis in Atmospheric Dispersion Modeling. *Pure Appl. Geophys.*, 162, 1893–1917.

20. Cox, R. T. (1946). Probability, Frequency, and Reasonable Expectation. *Am. J. Phys.*, 14, 1–13.

21. Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, UK.

22. Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice*, Chapman and Hall, CRC Press, Boca Raton, Florida.

23. Gelman, A., Carlin, J., Stern, H. and Rubin, D. (2003). *Bayesian Data Anaysis*, Chapman and Hall, CRC Press, Boca Raton, Florida.

24. Green, P. J. (1995). Reversible–Jump Markov chain Monte Carlo Computation and Bayesian Model Determination. *Biometrika* 82, 711–732.

25. Geyer, C. J. (1991). Markov chain Monte Carlo Maximum Likelihood, In *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, E.M. Keramidas (ed.), pp. 156–163.

26. Yee, E., Kosteniuk, P. R., Chandler, G. M., Biltoft, C. A. and Bowers, J. F. (1993). Statistical Characteristics of Concentration Fluctuations in Dispersing Plumes in the Atmospheric Surface Layer. *Boundary-Layer Meteorology* 65, 69–109.

27. Wilson, J. D., Flesch, T. K. and Harper, L. A. (2001). Micro-meteorological Methods for Estimating Surface Exchange with a Disturbed Windflow. *Agricultural and Forest Meteorology*, 107, 207–225.

28. Stull, R. B. (1988). *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

29. Rodean, H. C. (1996). *Stochastic Lagrangian Models of Turbulent Diffusion*, Meteorological Monographs, Vol. 26, No. 48, American Meteorological Society, Boston, MA.

TABLE 1

The mean, standard deviation (S.D.), and lower and upper bounds of the 95% highest posterior density (HPD) interval of the parameters $x_{s,k}$, $y_{s,k}$ and $q_{s,k}$ $(k = 1, 2)$ calculated from the samples for each cluster in Figures 3, 6 and 7 for $N_s = 2$, 3 and 4, respectively. These are the parameters that characterize the two source objects $(k = 1, 2)$ associated with the two clusters of samples.

	Parameter	Mean	S.D.	95% HPD
$k = 1$				
	x_s (m)	-254	11	$(-276, -234)$
	y_s (m)	-0.13	0.37	$(-0.85, 0.58)$
	q_s (g s^{-1})	1.04	0.05	(0.95,1.14)
$k = 2$				
	x_s (m)	-50	0.2	$(-50.5, -49.6)$
	y_s (m)	-0.00195	0.036	$(-0.07, 0.07)$
	q_s (g s^{-1})	1.02	0.01	(0.99,1.05)

TABLE 2

The mean, standard deviation (S.D.), and lower and upper bounds of the 95% highest posterior density (HPD) interval of the parameters $x_{s,k}$, $y_{s,k}$ and $q_{s,k}$ ($k = 1, 2, 3$) calculated from the samples for each cluster in Figures 12 and 13 for $N_s = 3$. These are the parameters that characterize the three source objects ($k = 1, 2, 3$) associated with the three clusters of samples.

	Parameter	Mean	S.D.	95% HPD
$k = 1$				
	x_s (m)	-288	16	$(-323, -260)$
	y_s (m)	20.8	1.1	$(18.4, 22.6)$
	q_s (g s^{-1})	0.956	0.094	(0.80,1.16)
$k = 2$				
	x_s (m)	-124	1.7	$(-127.5, -121.3)$
	y_s (m)	-20.1	0.2	$(-20.5, -19.7)$
	q_s (g s^{-1})	1.03	0.03	(0.97,1.09)
$k = 3$				
	x_s (m)	-50.1	0.2	$(-50.5, -49.6)$
	y_s (m)	-0.0073	0.039	$(-0.084, 0.067)$
	q_s (g s^{-1})	1.03	0.02	(0.99,1.05)

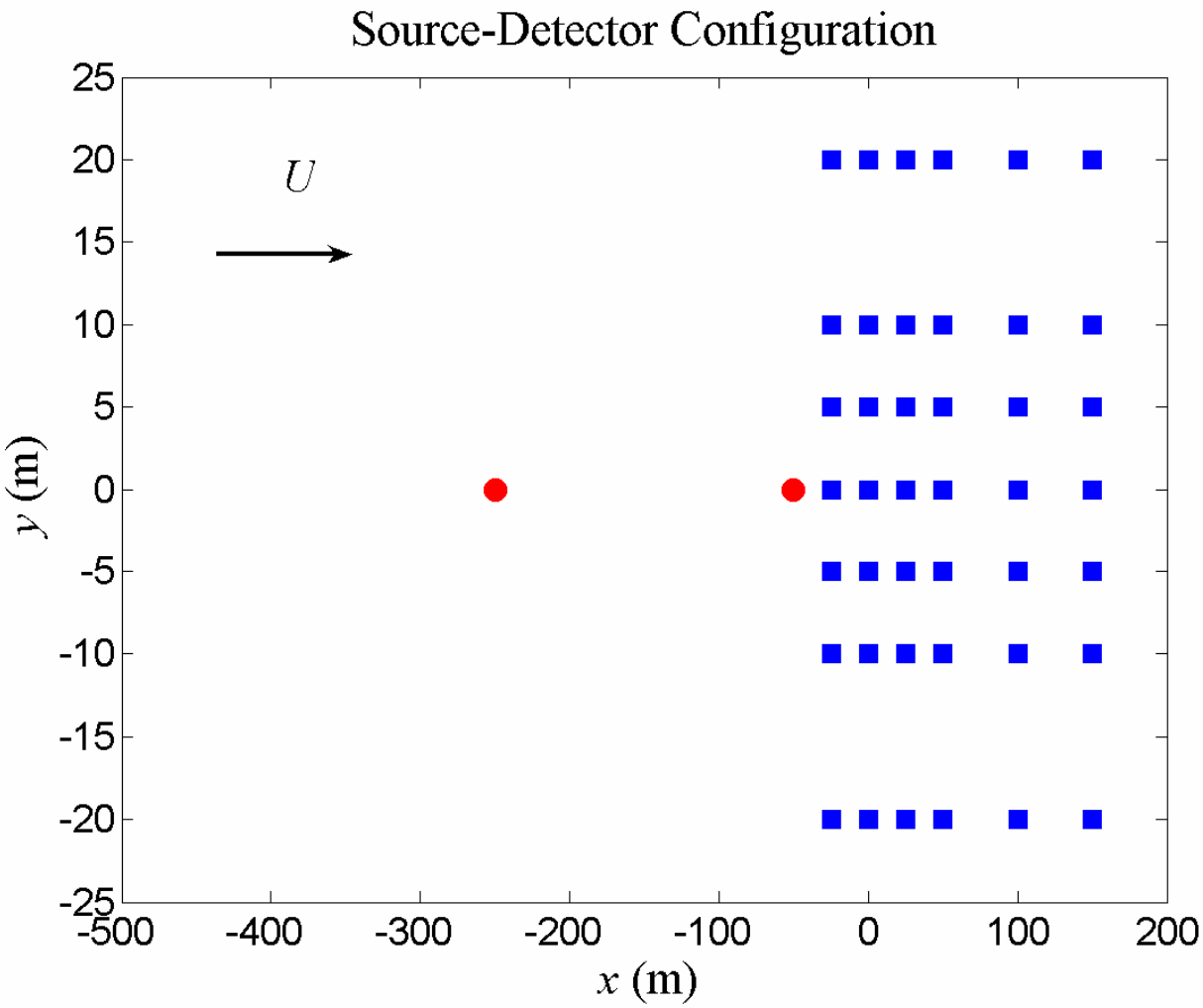

Figure 1. Two continuous point sources located upwind of an array of 42 detectors arranged as shown. A solid dot denotes a source and a solid square denotes a sensor.

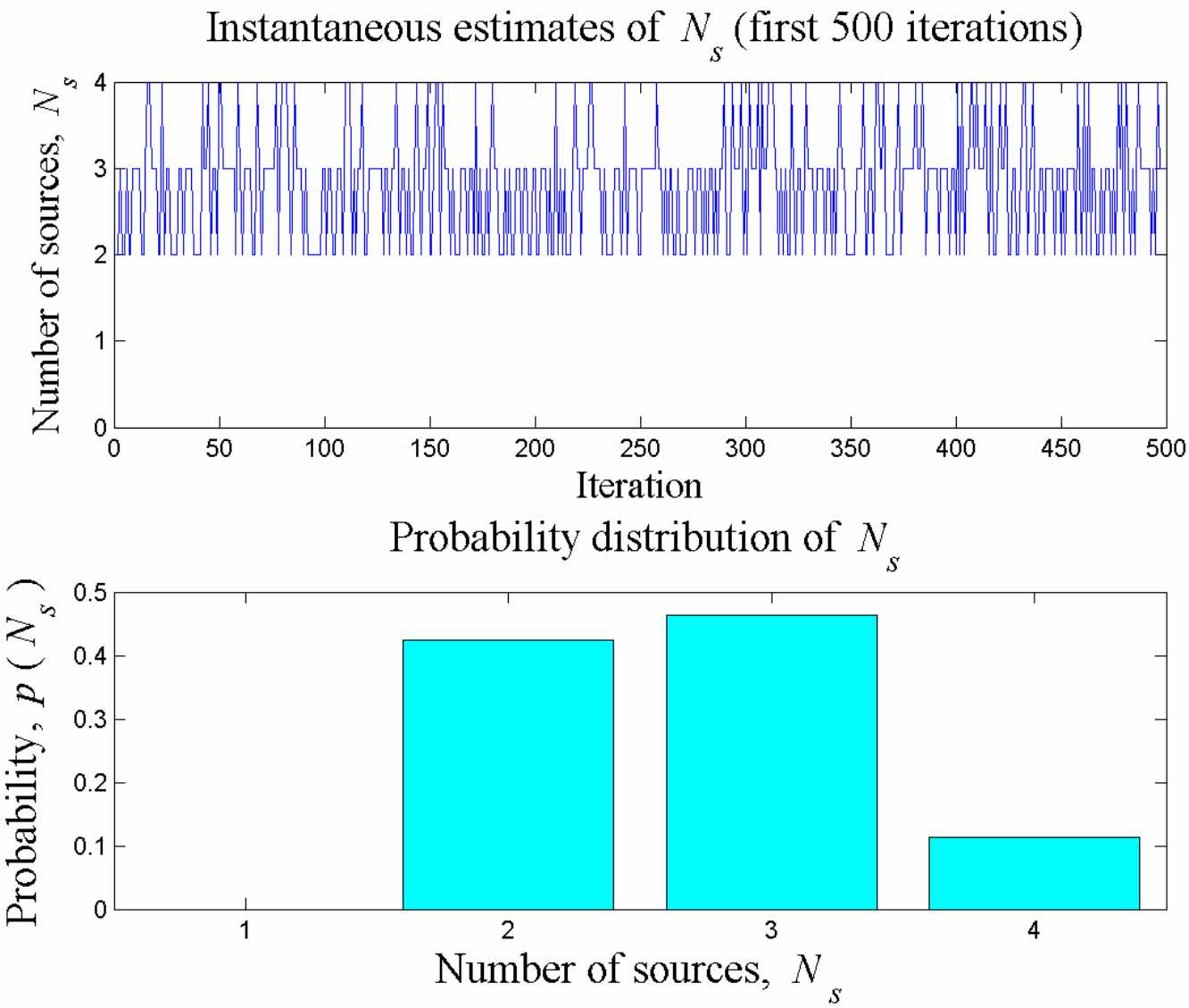

Figure 2. Instantaneous estimates of N_S for the first 500 post burn-in iterations of the reversible jump MCMC sampler (top) and the posterior probability distribution for the number of sources (bottom) estimated using the 50,000 post burn-in samples.

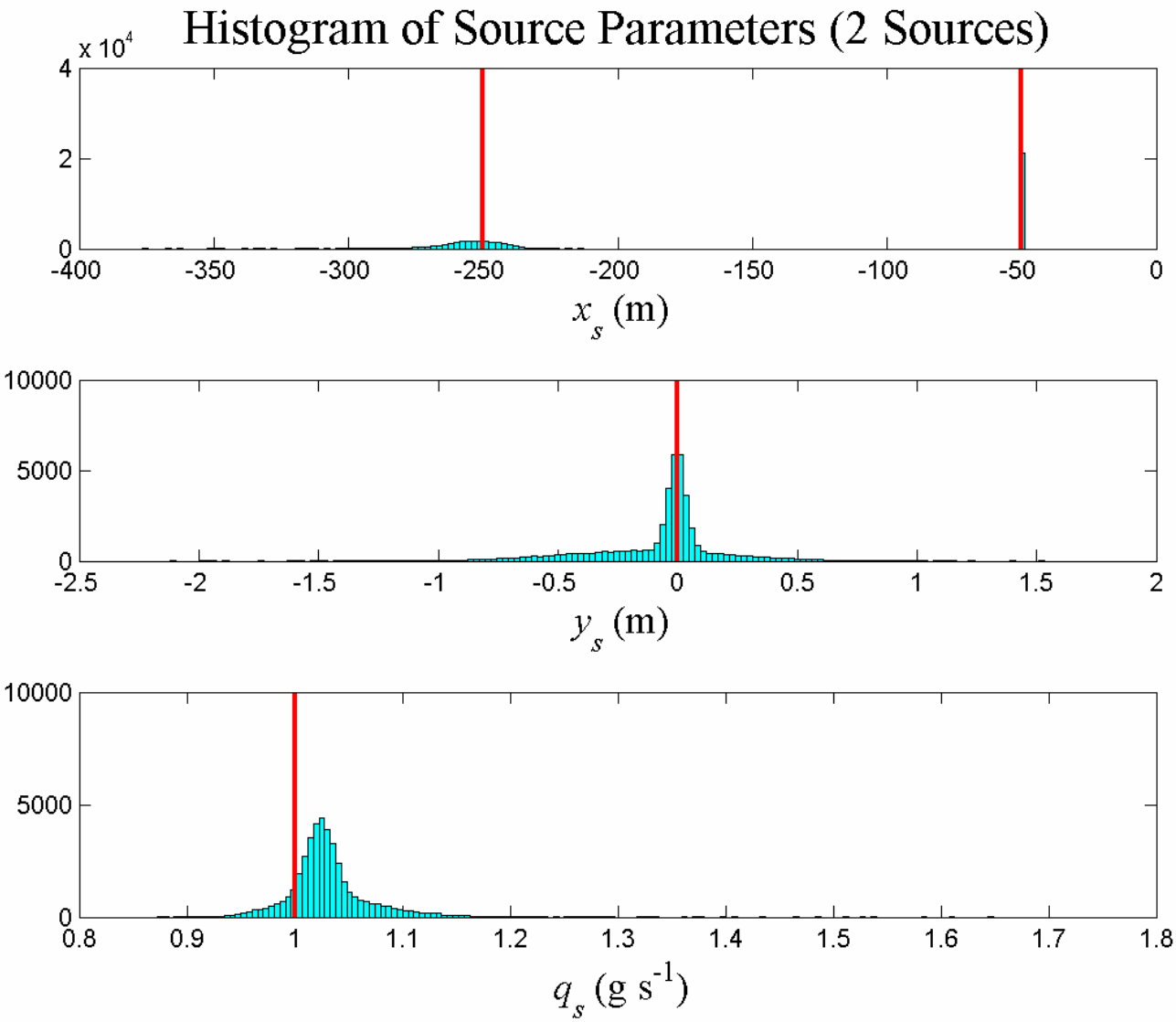

Figure 3. Histograms of the alongwind location (top panel), crosswind location (middle panel), and emission rate (bottom panel) of the two sources obtained from post burn-in samples with N_S = 2. The vertical line(s) in each panel mark the true values for the source parameters.

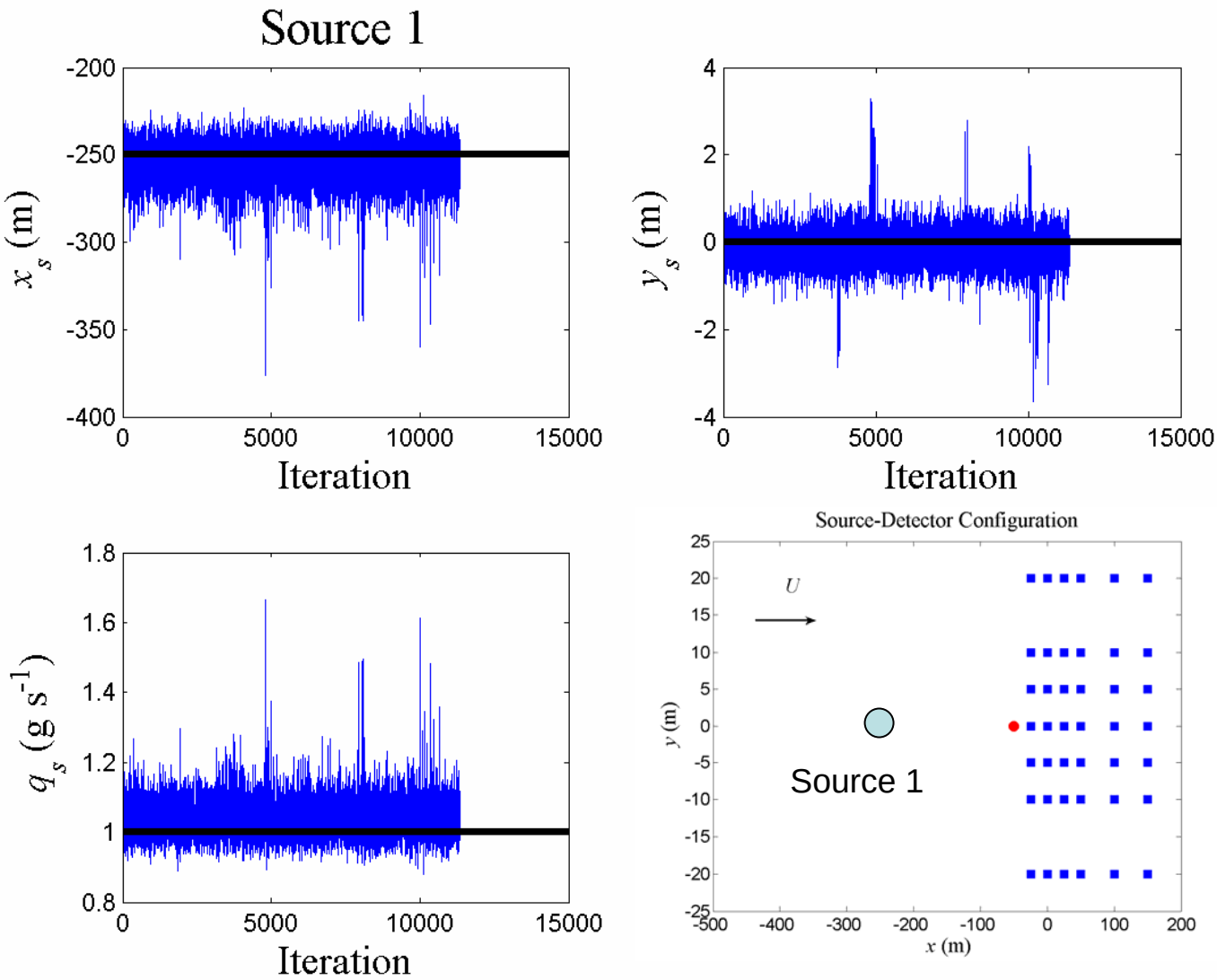

Figure 4. Example of a trace of source parameters for a subset of 12,000 samples extracted from the samples associated with the mode at x_S = -250 m in the histogram of x_S exhibited in Figure 3 (top panel). The solid horizontal lines are the true values for the source parameters.

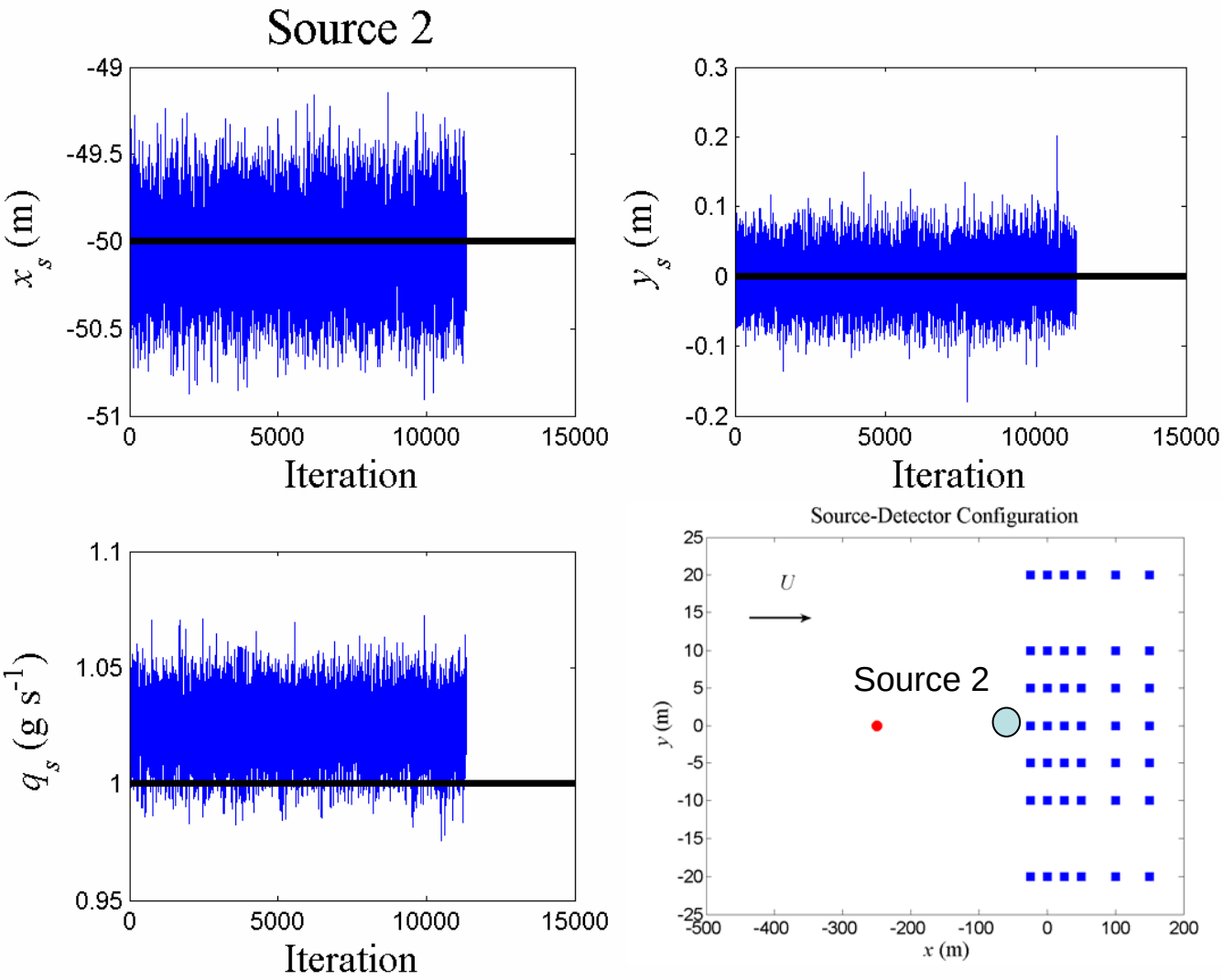

Figure 5. Example of a trace of source parameters for a subset of 12,000 samples extracted from the samples associated with the mode at x_S = -50 m in the histogram of x_S exhibited in Figure 3 (top panel). The solid horizontal lines are the true values for the source parameters.

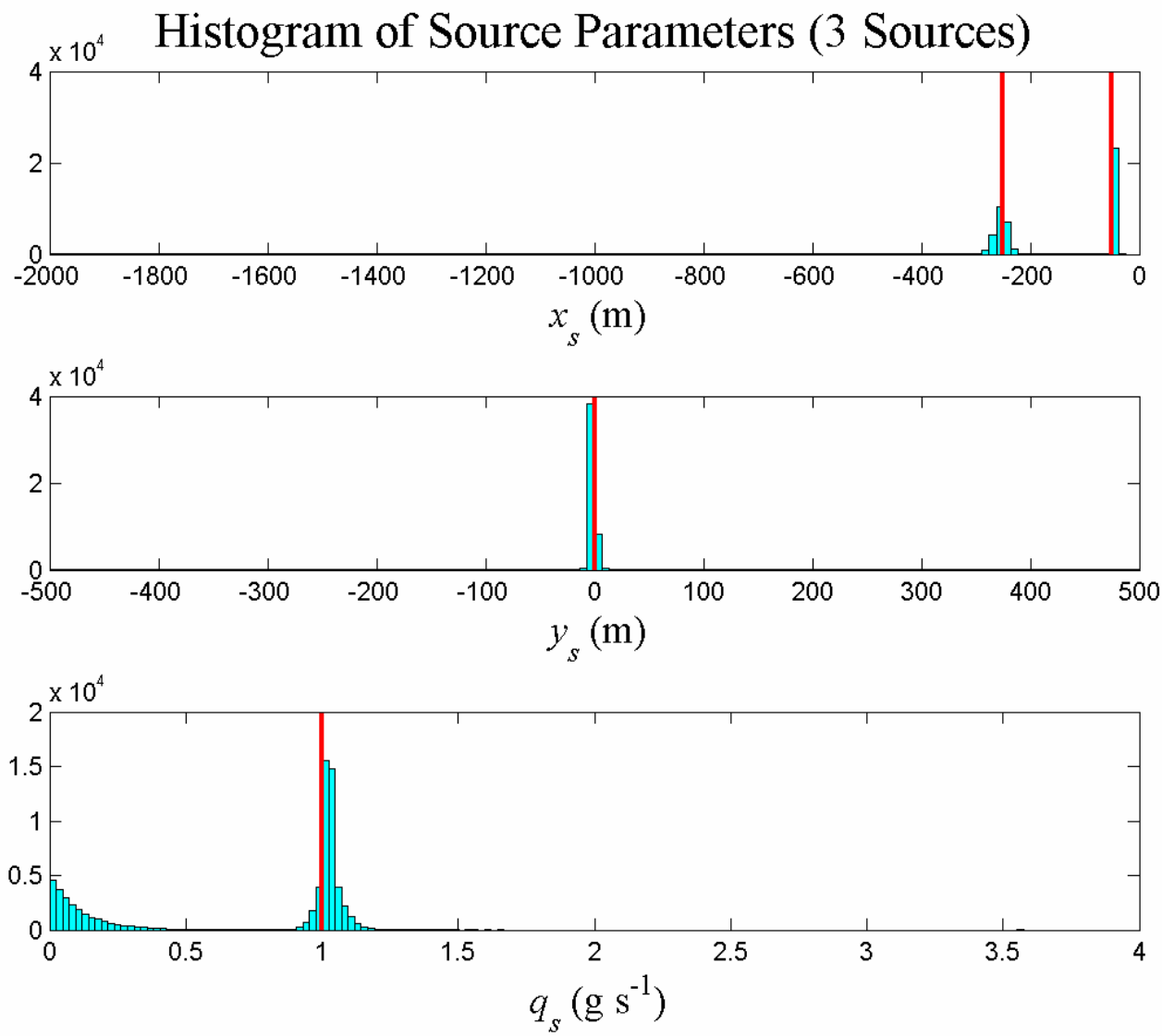

Figure 6. Histograms of the alongwind location (top panel), crosswind location (middle panel), and emission rate (bottom panel) of the two sources obtained from post burn-in samples with N_S = 3. The vertical line(s) in each panel mark the true values for the source parameters.

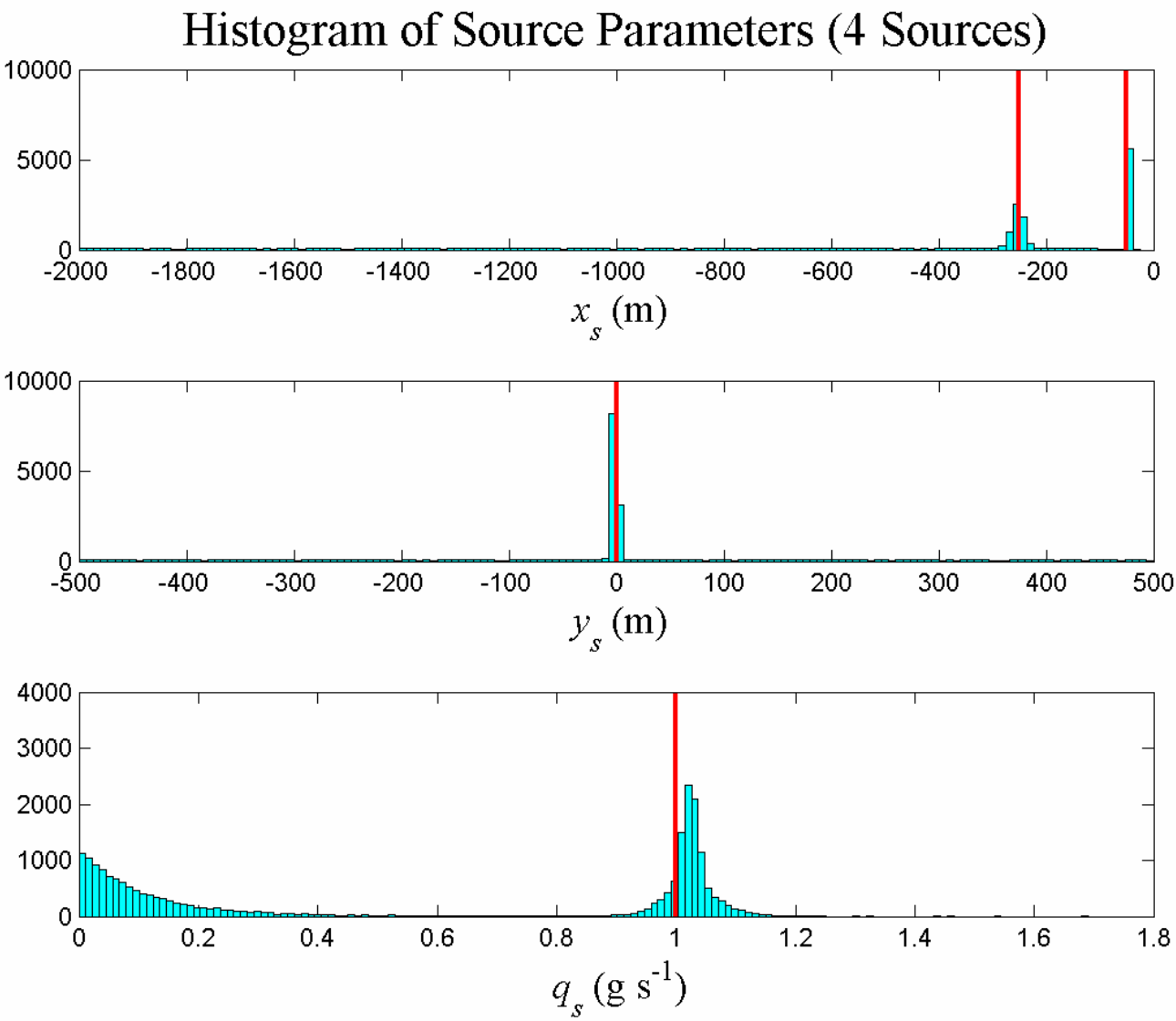

Figure 7. Histograms of the alongwind location (top panel), crosswind location (middle panel), and emission rate (bottom panel) of the two sources obtained from post burn-in samples with N_S = 4. The vertical line(s) in each panel mark the true values for the source parameters.

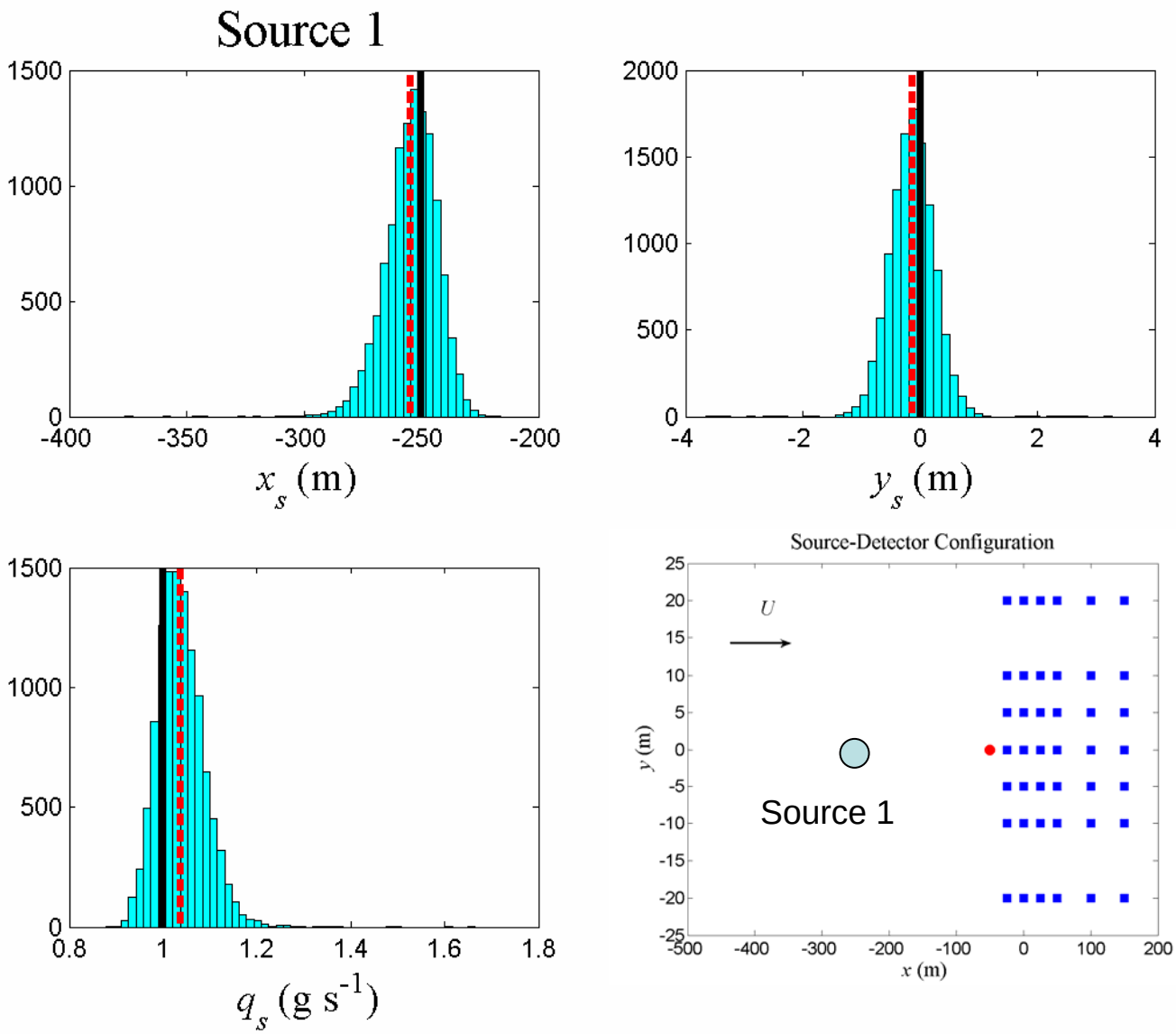

Figure 8. Histograms for the three parameters, namely alongwind location (top left frame), crosswind location (top right frame) and emission rate (bottom left frame) that characterize source 1 (bottom right frame). The solid vertical line indicates the true value of the parameter and the dashed vertical line corresponds to the best estimate of the parameter obtained as the posterior mean of the associated marginal posterior distribution.

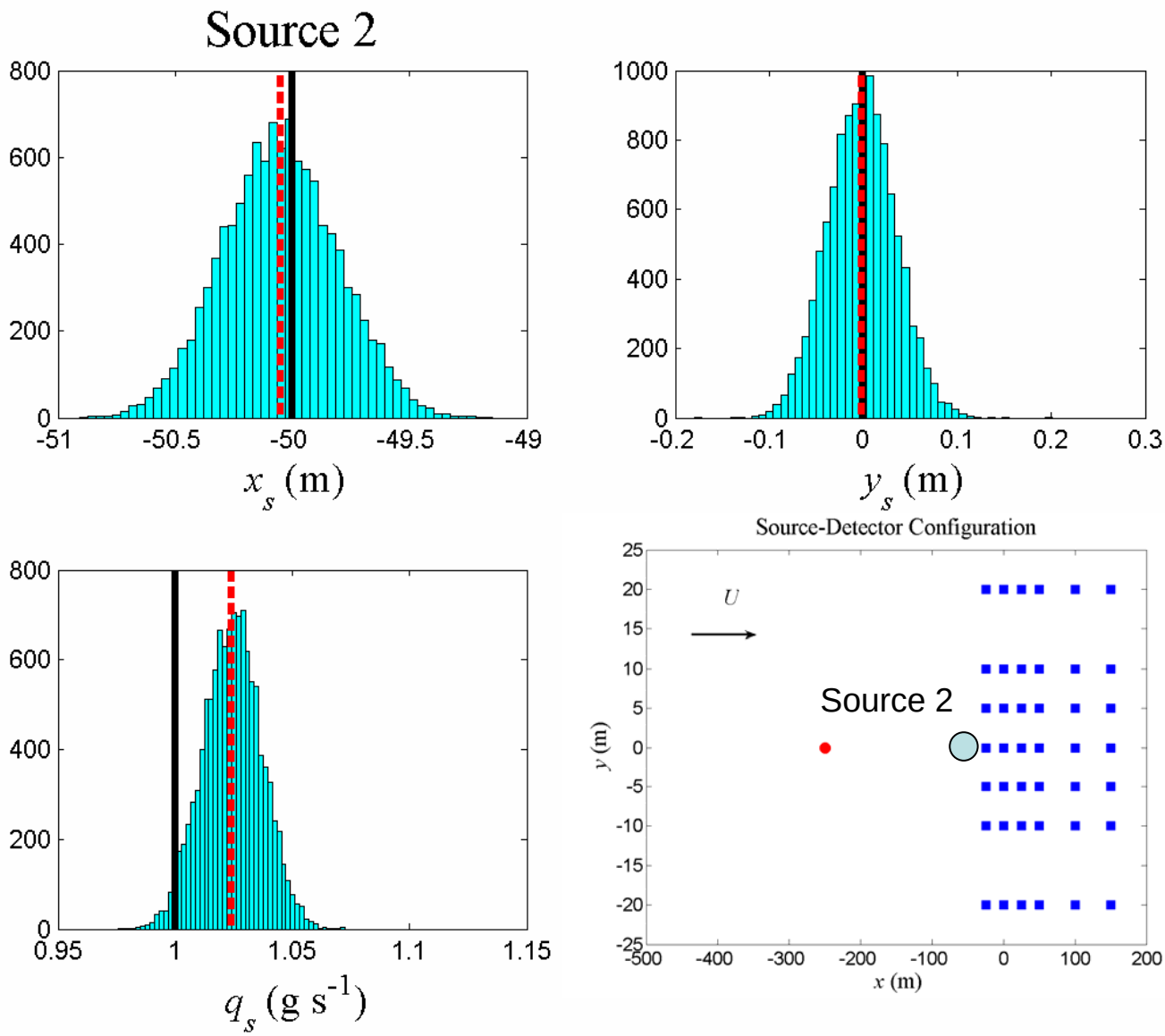

Figure 9. Histograms for the three parameters, namely alongwind location (top left frame), crosswind location (top right frame) and emission rate (bottom left frame) that characterize source 2 (bottom right frame). The solid vertical line indicates the true value of the parameter and the dashed vertical line corresponds to the best estimate of the parameter obtained as the posterior mean of the associated marginal posterior distribution.

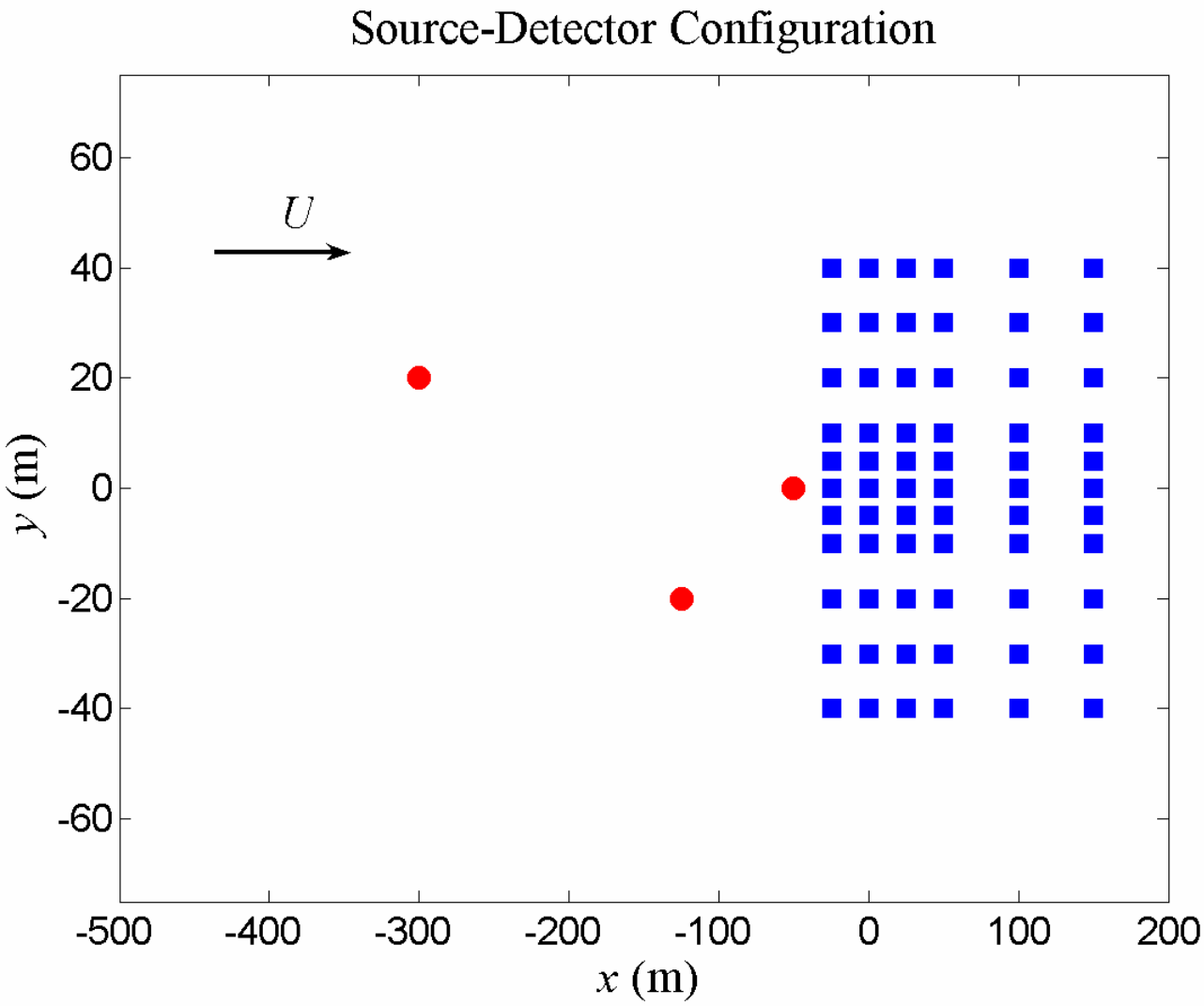

Figure 10. Three continuous point sources located upwind of an array of 66 detectors arranged as shown. A solid dot denotes a source and a solid square denotes a sensor.

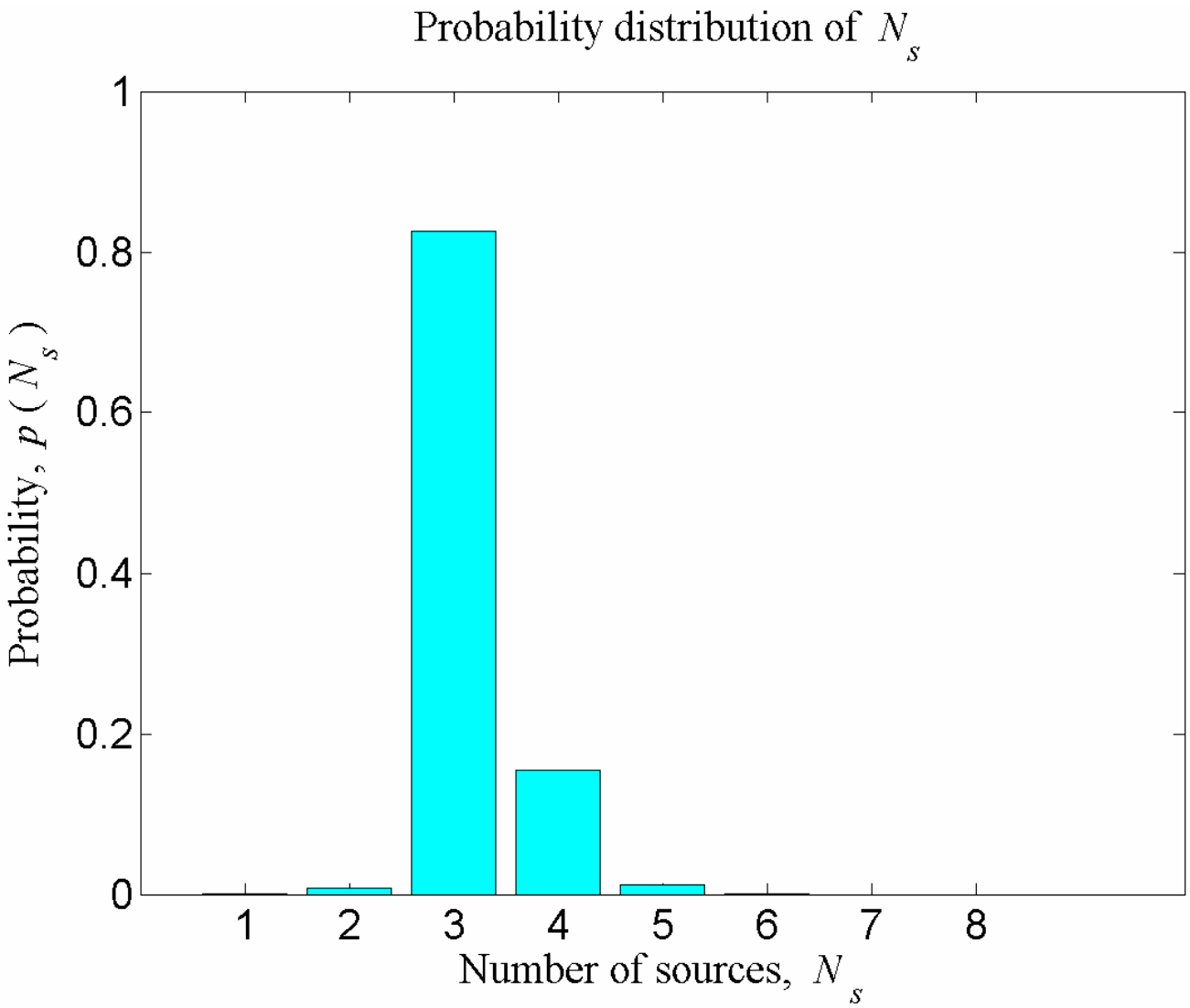

Figure 11. The posterior distribution for the number of sources estimated using the 50,000 post burn-in samples.

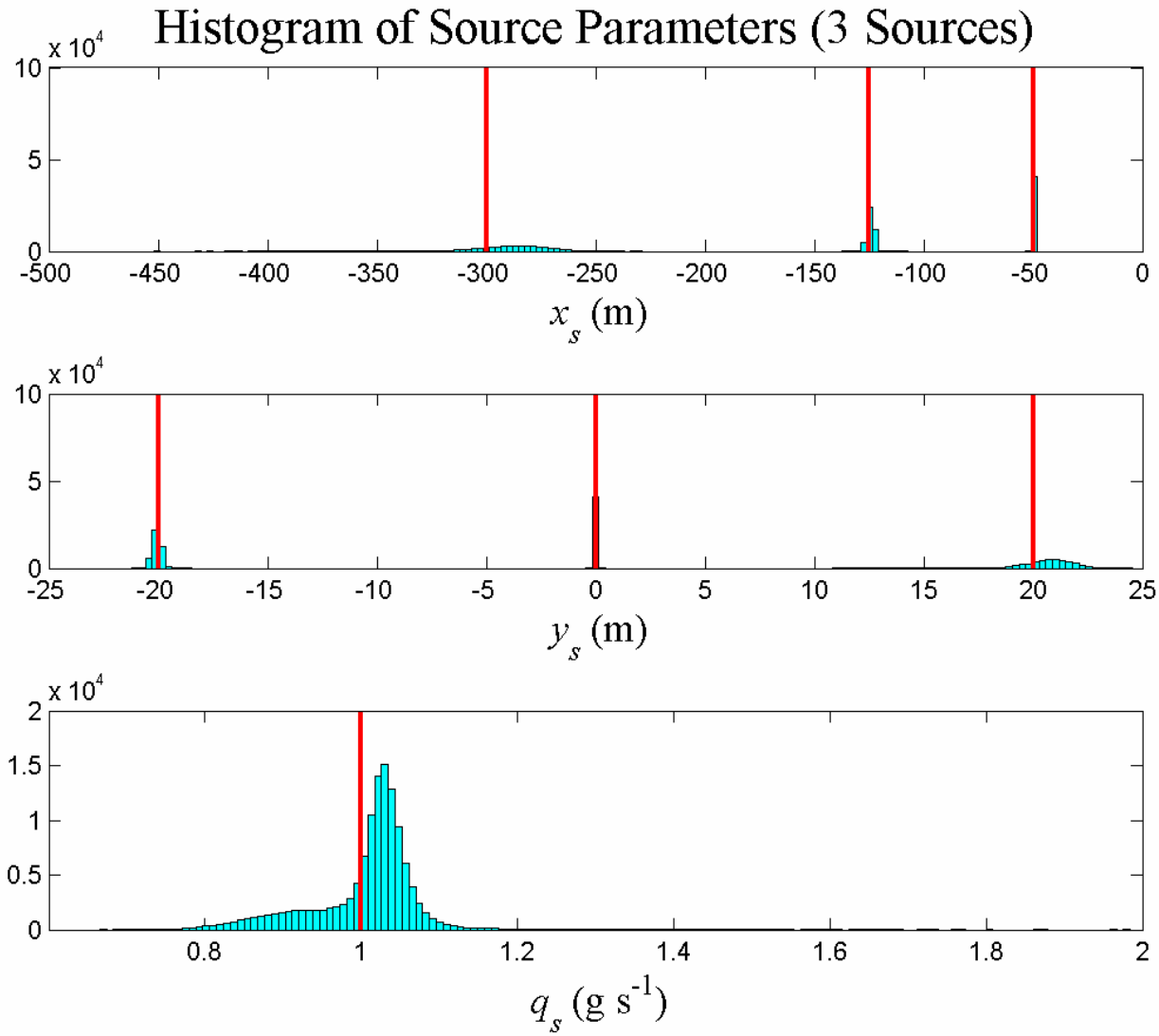

Figure 12. Histograms of the alongwind location (top panel), crosswind location (middle panel), and emission rate (bottom panel) of the three sources obtained from post burn-in samples with N_S = 3. The vertical line(s) in each panel mark the true values for the source parameters.

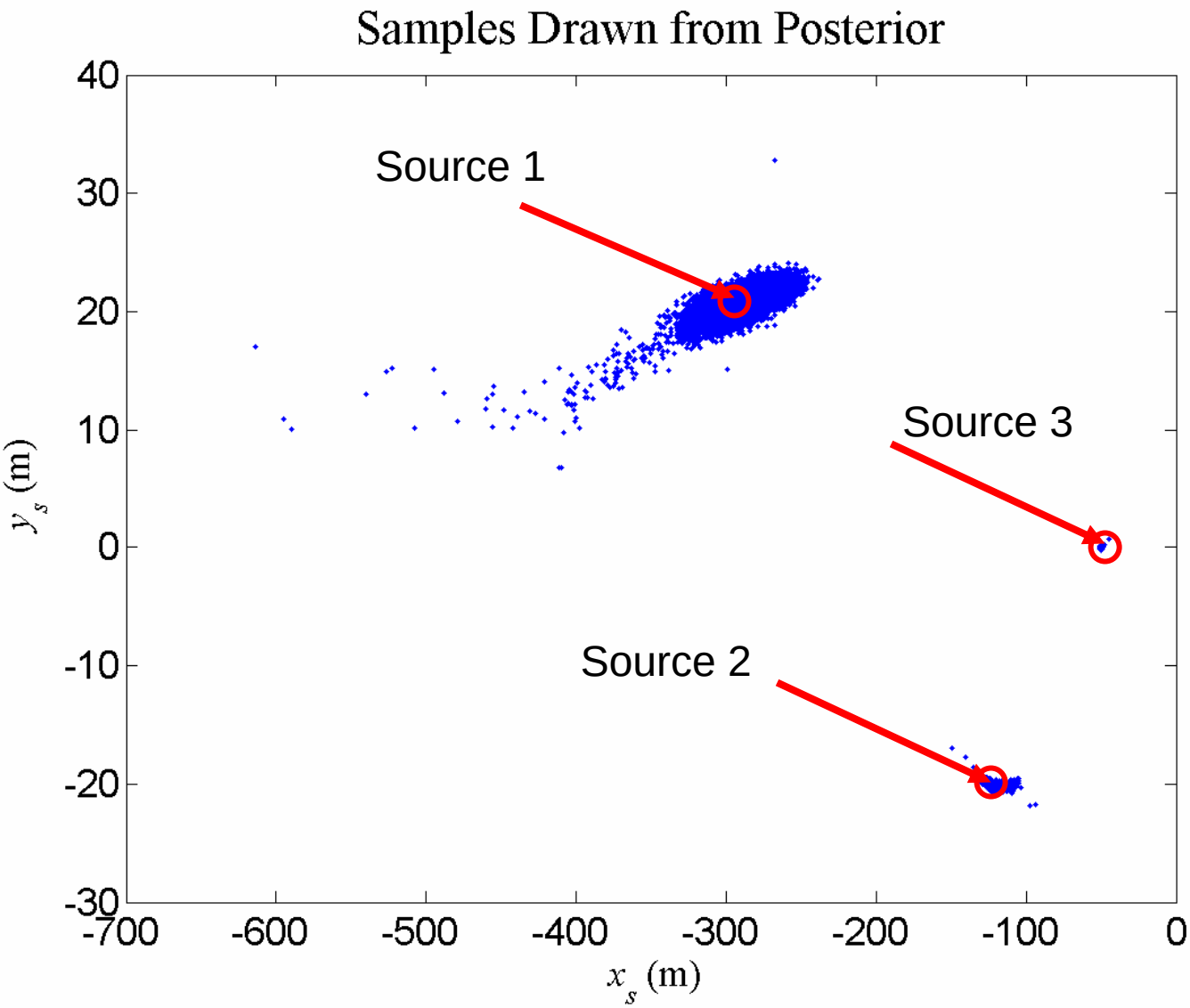

Figure 13. The post burn-in samples obtained for N_S = 3 projected onto the (x_S,y_S) subspace. The true alongwind and crosswind locations of the three sources are indicated by the open circles.

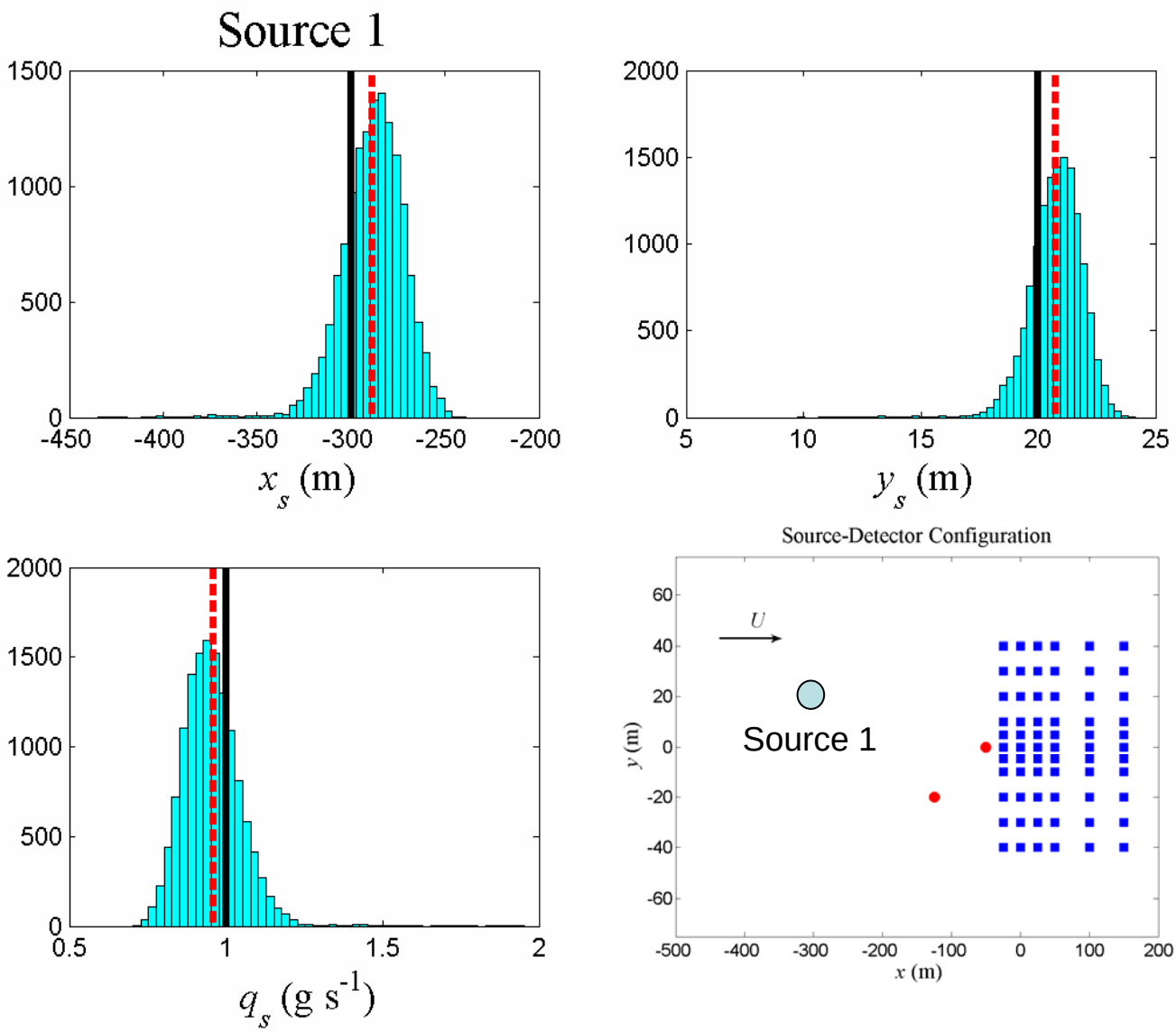

Figure 14. Histograms for the three parameters, namely alongwind location (top left frame), crosswind location (top right frame) and emission rate (bottom left frame) that characterize source 1 (bottom right frame). The solid vertical line indicates the true value of the parameter and the dashed vertical line corresponds to the best estimate of the parameter obtained as the posterior mean of the associated marginal posterior distribution.

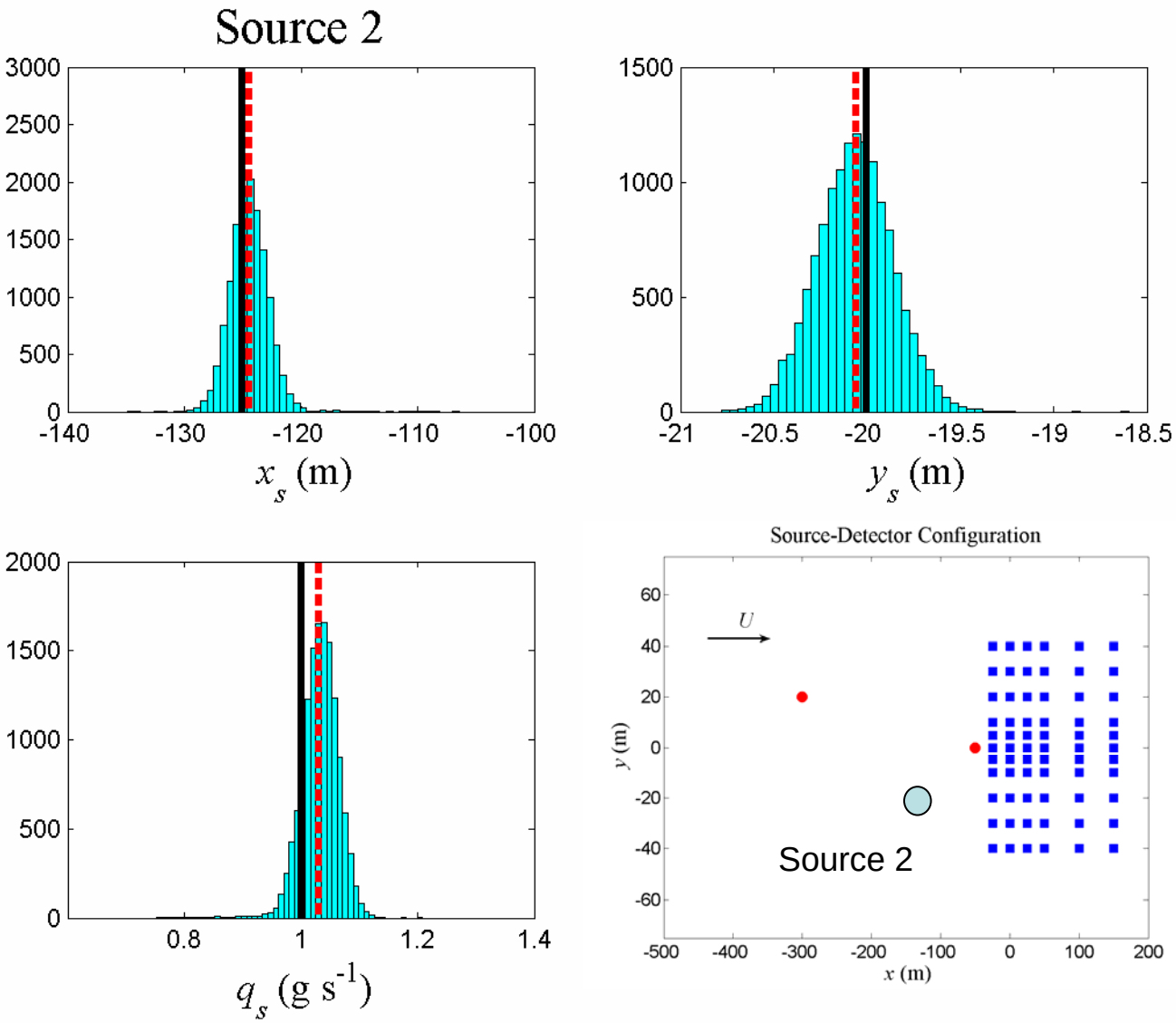

Figure 15. Histograms for the three parameters, namely alongwind location (top left frame), crosswind location (top right frame) and emission rate (bottom left frame) that characterize source 2 (bottom right frame). The solid vertical line indicates the true value of the parameter and the dashed vertical line corresponds to the best estimate of the parameter obtained as the posterior mean of the associated marginal posterior distribution.

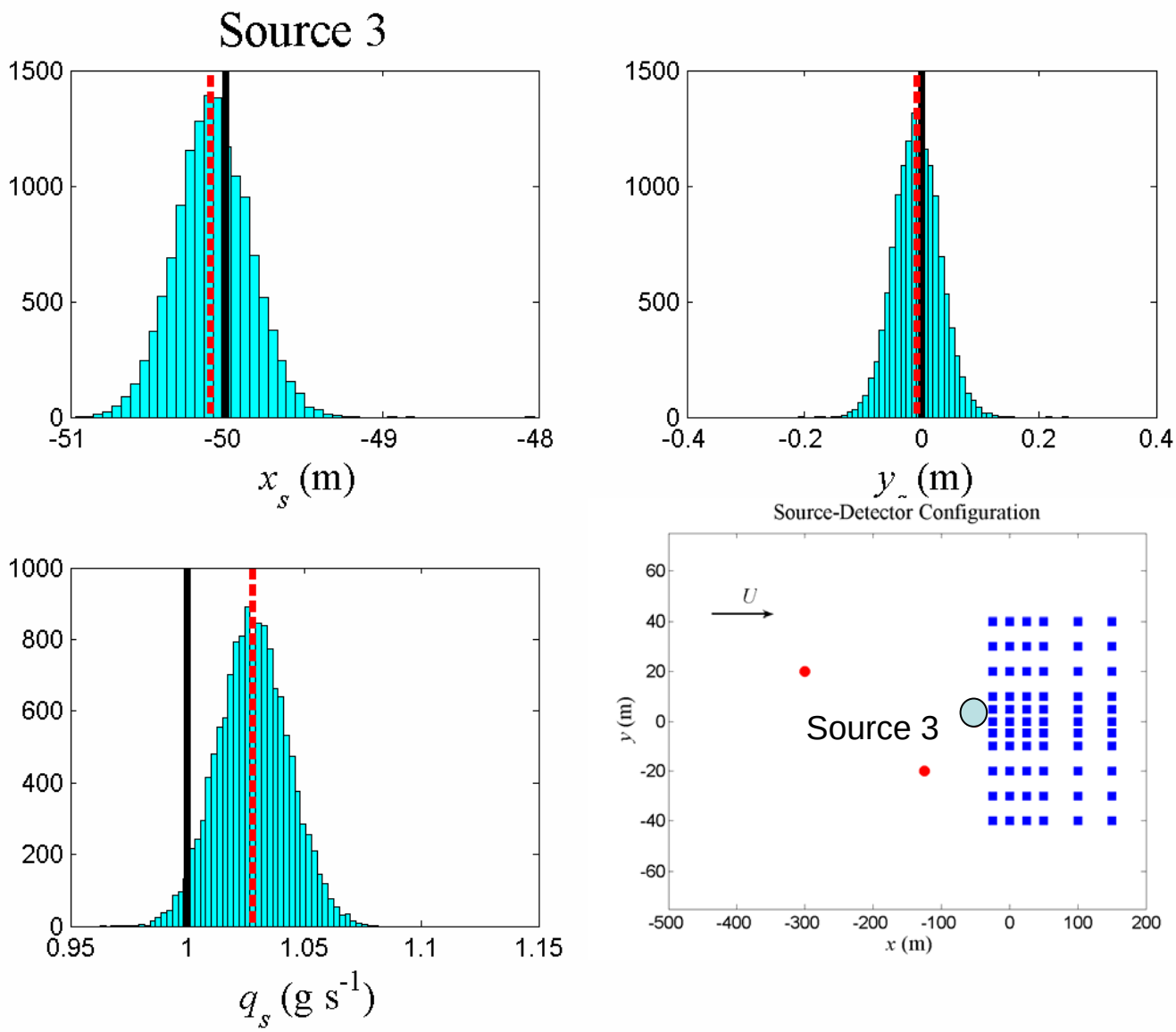

Figure 16. Histograms for the three parameters, namely alongwind location (top left frame), crosswind location (top right frame) and emission rate (bottom left frame) that characterize source 3 (bottom right frame). The solid vertical line indicates the true value of the parameter and the dashed vertical line corresponds to the best estimate of the parameter obtained as the posterior mean of the associated marginal posterior distribution.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for who the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in Section 8.)

   Defence R&D Canada – Suffield
   PO Box 4000, Station Main
   Medicine Hat, AB T1A 8K6

2. SECURITY CLASSIFICATION
   (overall security classification of the document, including special warning terms if applicable)

   Unclassified

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title).

   Bayesian Inversion of Concentration Data for an Unknown Number of Contaminant Sources (U)

4. AUTHORS (Last name, first name, middle initial. If military, show rank, e.g. Doe, Maj. John E.)

   Yee, Eugene

5. DATE OF PUBLICATION (month and year of publication of document)

   June 2007

6a. NO. OF PAGES (total containing information, include Annexes, Appendices, etc) 54

6b. NO. OF REFS (total cited in document) 29

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

   Technical Report, Final, January 1, 2006 – May 31, 2007

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

   CBRN Research and Technology Initiative

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)

   CRTI Project No. 02-0093RD

9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

   DRDC Suffield TR 2007-085

10b. OTHER DOCUMENT NOs. (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

   (x) Unlimited distribution
   ( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
   ( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
   ( ) Distribution limited to government departments and agencies; further distribution only as approved
   ( ) Distribution limited to defence departments; further distribution only as approved
   ( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally corresponded to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected).

   Unlimited

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C) or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

In this paper, we address the inverse problem of source reconstruction for the difficult case of multiple sources when the number of sources is unknown *a priori*. The problem is solved using a Bayesian probabilistic inferential framework in which Bayesian probability theory is used to derive the posterior probability density function for the number of sources and for the parameters (e.g., location, emission rate, release duration) that characterize each source. A mapping (or, source–receptor relationship) that relates a multiple source distribution to the concentration data measured by the array of sensors is formulated based on a forward-time Lagrangian stochastic model. A computationally efficient methodology for determination of the likelihood function for the problem, based on an adjoint representation of the source–receptor relationship and realized in terms of a backward-time Lagrangian stochastic model, is described. An efficient computational algorithm based on a reversible jump Markov chain Monte Carlo (MCMC) method is formulated and implemented to draw samples from the posterior density function of the source parameters. This methodology allows the MCMC method to jump between the hypothesis spaces corresponding to different numbers of sources in the source distribution and, thereby, allows a sample from the full joint posterior distribution of the number of sources and source parameters to be obtained. The proposed methodology for source reconstruction is tested using synthetic concentration data generated for cases involving two and three unknown sources.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifies, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Source reconstruction, Multiple sources, Bayesian inference, Markov chain Monte Carlo, Lagrangian stochastic models, adjoint systems